AD-A047 667 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCHO--ETC F/G 21/5

DEVELOPMENT OF COST ESTIMATING RELATIONSHIPS FOR AIRCRAFT JET C--ETC(U)

SEP 77 R A BREGLIO, R F WRIGHT

UNCLASSIFIED AFIT-LSSR-31-77B NL

1 OF 2

AD
A047667

3

UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

3

DEVELOPMENT OF COST ESTIMATING
RELATIONSHIPS FOR AIRCRAFT JET
CORE-ENGINE OVERHAUL COSTS

Robert A. Breglio, Jr., LCDR, USN
Richard F. Wright, Lieutenant, USN

LSSR 31-77B

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deliterious information are contained therein. Furthermore, the views expressed in the document are those of the author and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.

USAF SCN 75-20B AFIT Control Number LSSR 31-77B

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. Please return completed questionnaires to: AFIT/SLGR (Thesis Feedback), Wright-Patterson AFB, Ohio 45433.

1. Did this research contribute to a current Air Force project?

   a. Yes b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?

   a. Yes b. No

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency received by virtue of AFIT performing the research. Can you estimate what this research would have cost if it had been accomplished under contract or if it had been done in-house in terms of manpower and/or dollars?

   a. Man-years ________ $ ________ (Contract).

   b. Man-years ________ $ ________ (In-house).

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3 above), what is your estimate of its significance?

   a. Highly Significant b. Significant c. Slightly Significant d. Of No Significance

5. Comments:

________________________ ________________________
Name and Grade Position

________________________ ________________________
Organization Location

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

14

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
AFIT-LSSR-31-77B

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
6 DEVELOPMENT OF COST ESTIMATING RELATIONSHIPS FOR AIRCRAFT JET CORE-ENGINE OVERHAUL COSTS.

5. TYPE OF REPORT & PERIOD COVERED
9 Master's Thesis,

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
10 Robert A. Breglio, Jr., Lieutenant Commander, ~~USN~~
Richard F. Wright, Lieutenant, USN

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Graduate Education Division
School of Systems and Logistics
Air Force Institute of Technology, WPAFB, OH

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Department of Research and Administrative Management
AFIT/LSGR, WPAFB, OH 45433

12. REPORT DATE
11 Sep[redacted]77

13. NUMBER OF PAGES
146 12 160 P

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Ø12 25Ø

18. SUPPLEMENTARY NOTES

APPROVED FOR PUBLIC RELEASE AFR 190-17.

JERRAL F. GUESS, CAPT, USAF
Director of Information

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cost Estimating
Aircraft
Jet Engine

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Thesis Chairman: Edward Karnasiewicz, Major, USAF

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Cost estimation is a wide open area within the Department of Defense and accurate cost estimating models are a valuable tool in the life cycle costing of a weapon system.

This research effort utilized multiple linear regression analysis to develop parametric cost models or cost estimating relationships (CERs) for jet engine depot overhaul costs. Both engine operating parameters, e.g., turbine inlet temperature, RPM, etc., and engine physical characteristics, e.g., length, weight, etc., were considered as probable cost drivers. Extensive analysis was performed to determine the reliability of the data base. The major finding of this study was that models can be developed with acceptable explanatory power, with respect to variation in the data base, using data of questionable reliability. Thus, a model developed should not be accepted on the basis of explanatory power alone but should be tested further to determine its utility as a cost estimator.

LSSR 31-77B

DEVELOPMENT OF COST ESTIMATING RELATIONSHIPS FOR AIRCRAFT JET CORE-ENGINE OVERHAUL COSTS

A Thesis

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Facilities Management

By

Robert A. Breglio, Jr., BS
Lieutenant Commander, USN

Richard F. Wright, BS
Lieutenant, USN

September 1977

Approved for public release;
distribution unlimited

This thesis, written by

Lieutenant Commander Robert A. Breglio, Jr.

and

Lieutenant Richard F. Wright

has been accepted by the undersigned on behalf of the faculty of the School of Systems and Logistics in partial fulfillment of the requirements for the degree or

MASTER OF SCIENCE IN LOGISTICS MANAGEMENT

DATE: 7 September 1977

COMMITTEE CHAIRMAN

READER

ACKNOWLEDGMENTS

We express our sincere appreciation to the many individuals who provided invaluable assistance in this research. We thank our advisor, Major Edward Karnasiewicz for his guidance and patience and our reader, Captain Joel B. Knowles for his constructive comments. We especially thank Captain Frederick P. Lawrence who made a significant contribution to this research by providing his invaluable statistical expertise and a great deal of his time. Further we extend gratitude to Mr. Mike Barga and Richard McNally of AFAPL/TBP and Mr. Roger Steinlage of AFLC/ACRCC for their continued support and assistance.

We express sincere appreciation to our typist, Joyce Clark, whose superior skill and knowledge were of unmeasurable value to us.

However, the most significant contribution was provided through the patience, understanding and support of Bob's wife, Donna, and Sharron Bowser.

TABLE OF CONTENTS

Chapter | Page

Chapter | Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

United States Air Force (USAF) acquisition and logistics efforts in support of aircraft turbine engines constitute a major management and fiscal program. For example, the active inventory of USAF aircraft turbine engines is approximately 38,000 which represent a total investment of some $10 billion in 1975 dollars. Additionally, nearly $500 million are spent annually to support and maintain this active inventory of jet engines (24:1).

During the past decade the annual cost to operate and support (O&S) USAF systems (including aircraft jet engines) has increased severalfold (24:1). An unquantifiable portion of the rise in O&S costs has been attributed to the lack of emphasis placed on downstream operating costs during the early conceptual and development phases of a jet engine's life cycle (2:1).

The USAF has developed a program for providing cost estimates. Air Force Regulation 173-1 delineates major command responsibilities to accomplish the task of cost estimating for proposed and developing systems. In accordance with this directive, USAF cost estimates for aircraft jet engines are developed by the

major command responsible for a particular phase of an engine's life cycle (22:3). The Headquarters, Air Force Logistics Command (AFLC) is responsible for providing preliminary cost estimates on proposed and developing jet engines which are modifications of existing jet engines. The AFLC is responsible for estimating those costs in support of jet engines which will be funded by the AFLC once full-scale development is complete. The primary area of concern to the AFLC is O&S costs which encompasses investment, maintenance, and support costs, including depot level overhaul cost. Within the AFLC, the Independent Cost Analysis Division (ICAD) is the responsible agency for developing required cost estimates (23).

Headquarters, Air Force Systems Command (AFSC), on the other hand, is responsible for providing preliminary cost estimates for jet engines which must undergo a research or development effort. Within the AFSC, the Air Force Aero Propulsion Laboratory (AFAPL) is the focal point for jet engine development (24:29-35) and, as such, is responsible for coordinating life cycle costs (LCC) estimates

> . . . during system design, concept evolution and selection, [and] fullscale development . . . to ensure [that] appropriate trade-offs [are made] among investment costs, ownership costs, . . . and performance [7:5].

PROBLEM STATEMENT

One area for USAF jet engine O&S cost control is depot level overhaul cost estimates for proposed USAF jet engines (23:i; 12).

It is, however, the opinion of ICAD and AFAPL personnel interviewed that existing cost estimating relationships (CERs) for predicting jet engine overhaul costs do not provide accurate enough estimates (16; 18; 13:74).

BACKGROUND

The U.S. budget allocation for defense spending has become more austere each year. Between fiscal year 1964 and fiscal year 1976, the allotment for defense has been reduced from 43 percent to 25 percent of the Federal budget and actual purchasing power is less now than 12 years ago (6:10-76).

While U.S. defense spending has been decreasing, the cost of operating and supporting weapon systems within the Department of Defense (DOD) has been increasing (12:1). This situation presents a significant problem for the DOD that has received considerable attention in recent years with respect to finding a solution.

In February, 1976, Deputy Secretary of Defense W. P. Clements, Jr. released a memorandum to the Secretaries of the Military Departments wherein he said,

> I am seriously concerned with the continuing growth of the fraction of the total DOD resources needed to operate and support our weapons and the decline in funds for new weapon procurement . . . [We must] increase emphasis on controlling the outyear operating and support costs of weapon systems during the development and acquisition phase both through attention to design, procurement, and

> support planning . . . I am equally concerned that insufficient attention is being paid to controlling eventual system O&S [operating and support] costs during conceptual, validation and full-scale development phases of new systems. My objective is to achieve an overall reduction in the fraction of each Service's outyear budget allocated to O&S costs of the new systems we are developing . . . From this time, each DSARC [Defense System Acquisition Review Council] review is to specifically address the O&S cost impact of new systems compared to those to be replaced or augmented . . . [21:A1-2].

The DOD, concerned with reducing the fraction of each Service's outyear budget allocated to O&S costs of new weapons, is utilizing the Design-to-Cost (DTC) program and the requirement to develop life cycle costs (LCC) estimates for developing weapon systems to: (a) achieve the lowest possible unit cost, and (b) to provide DOD decision makers with a trade-off alternative based on the lowest estimated LCC (26; 27).

The DOD has stipulated that costs of acquisition and ownership are to be established as separate cost elements and translated into firm DTC and LCC requirements for a system selected for full-scale engineering development. Weapon system program actions are to be evaluated against these requirements with the same rigor as the evaluation of system technical specifications and requirements (28:9). Design-to-Cost requires the translation of acquisition and ownership (development, production, operation and support) costs into "design-to" requirements and the utilization of these requirements to evaluate system development. It also requires that trade-offs be made among

performance, development schedule and costs if the end result is to be production and O&S cost reduction (20:84). Life cycle cost includes all costs from the conceptual phase of a system's acquisition cycle through validation, full-scale development, production, deployment, and the revitalization and disposition phases (30).

To reduce the O&S costs, efforts are being expended to design in improved reliability and maintainability, and to use lessons learned from existing field systems in the design of new systems. Cost estimates are being used for source selection, contract negotiations and the establishment of contract incentives. Cost estimates are also used to provide O&S cost predictions early in the acquisition cycle so that considerations of design alternatives to maintenance and logistics concepts will be available to provide significant cost reductions (19:11).

COST ESTIMATING

A memorandum in December, 1971 from the then Secretary of Defense Packard, advised all Service Secretaries that parametric cost estimates and analyses for each weapon system in the acquisition cycle were to be incorporated in each Defense System Acquisition Review Council (DSARC) presentation. This memorandum was followed in January, 1972 by a memorandum from the then Secretary of Defense Laird which reaffirmed the importance of cost estimates in

the acquisition of weapon systems and established a Cost Analysis Improvement Group (CAIG) within the Office of the Secretary of Defense (OSD). The responsibility of the CAIG is to review cost estimates presented by Services to the DSARC and to develop uniform criteria to be used by all DOD units in making cost estimates (29: Encl 2).

The purpose of cost estimating is to produce reliable cost estimates for use in decision making during the acquisition process. Cost estimates are used by DOD decision-makers to weigh the potential values of a developing system against its predicted LCC. The same estimates are also used for (8:153-154):

1. Planning: Cost estimates provide information to support the long-range plans for proposed weapon systems and are also essential considerations in prediction of the LCC components of a developing system.

2. Budget preparation: Within the DOD, cost estimates are used to support both funding requests from the Congress and funding apportionment among the services.

3. Contract pricing: During procurement, government cost estimates are used to establish price negotiations objectives and to provide government contract negotiators with cost estimates independent of the contractors' estimates.

4. Estimating the cost of contract changes: Cost estimates

are useful in providing an insight to project managers of the impact program changes will have on the system's LCCs.

5. As a measurement tool for the control of programs: Cost estimates provide "yardsticks" to measure the predicted program costs versus the actual program costs.

The Air Force is utilizing cost estimating models to estimate both weapon system development costs and depot level overhaul costs which will be included as components of LCC estimates of developing weapon systems. Cost estimates are made by developing cost estimating relationships (CER) from cost data, physical characteristics, and operating parameters of existing weapon systems.

CERs are mathematical models that describe the cost of an item or of an activity as a function of one or more independent variables. To be useful, CERs should have the following characteristics (31:Ch. 3):

1. They should be relatively easy to use.

2. They should be relatively inexpensive so that the cost of development is held to a minimum.

3. They should have minimal data requirements since they are usually developed and tested during periods when little is known about the systems for which the CERs are being developed.

Although CERs are used to predict costs of new or developing systems, there are limitations to their use. For example, CERs

should generally not be applied to estimate costs of radically new systems since there will be little prior knowledge from which to make meaningful estimates. Additionally, since CERs are developed from minimal data, periodic adjustments may be required to allow for changes in economic trends, design, and maintenance policy (31:Ch. 3). Since each cost estimate has its own conceptual ground rules, assumptions, and criteria, each is prepared for a specific purpose at a particular time within a system's acquisition cycle. For these reasons, the development of an all-purpose CER may not be feasible (22:3).

Two basic types of cost estimating procedures used by the Air Force are (8:465):

1. the traditional industrial engineering approach to cost estimating which relies on detailed analysis of the system being developed, and
2. the parametric cost estimating procedure which makes use of historical cost data from existing systems that are considered to be similar in design and level of technology to the developing system.

THE INDUSTRIAL ENGINEERING METHOD

The industrial engineering approach to cost estimating relies on the availability of detailed information about the system under development. It requires the use of all the costs associated with

those elements of the developing system which, when combined, will form the final product (8:153-157). It has the advantage of being able to be tailored to a specific system and often to a specific contractor. Its major strength is the accuracy of the estimate it provides.

In order to utilize the industrial engineering method, certain facts about the developing system must be known: generally, the costs of direct and indirect labor and material which will be required to produce the system. Because of the level of detail required, this method is difficult and expensive to use (8:153-157). An additional drawback is that as cost estimates for various components of the weapon system pass through succeeding levels of management, the estimates run ". . . the risk of becoming inflated through failure to identify the [subjective] contributions of managers at each level of summation [8:157]."

THE PARAMETRIC METHOD

The parametric method of cost estimating relies upon functional relationships between the cost of the alternative and the various characteristics or parameters of the alternative. The results of a parametric estimate depend directly upon the analyst's ability to identify the cost influencing characteristics or parameters of alternatives and to use these characteristics or parameters to develop CERs.

A notable strength of the parametric method is that this approach can provide cost estimates during the conceptual stage of the acquisition process before detailed engineering plans are available. Parametric estimates predict cost by developing mathematical relationships among the variables, e.g., cost and performance characteristics and/or physical characteristics, of similar systems (17:3).

However, in utilizing the parametric method to develop cost estimates, there are difficulties which may be encountered. These include:

1. The system for which the estimate is being developed must be closely related in performance and design technology to the systems used to provide the estimating data (28:9).
2. The accuracy of the estimate developed is limited by the fact that physical and performance characteristics might not explain all of the variance in the predicted cost (28:9).
3. The cost estimate and the CER might become obsolete quickly under conditions of changing technology (8:156-157).

Aside from the major difficulties that may be encountered when utilizing the parametric method to develop CERs, there are advantages to the use of this approach:

1. As previously identified, cost estimation is possible early in the conceptual stage prior to detailed engineering

plans being available (8:157).

2. The use of past experience incorporates engineering and design specification changes in the data base which were not identifiable during conceptual development of the system (28:10).

3. Cost estimates can be made quickly once a CER has been developed (28:10).

RESEARCH OBJECTIVE AND HYPOTHESIS

Present USAF jet engine cost estimating relationships for predicting jet engine depot level overhaul costs, utilize total engine depot level overhaul cost as the dependent variable. However, other than the engine physical characteristic "engine dry weight" the independent variables used by the AFLC to develop CERs published in AFLC Pamphlet 173-4 (Appendix A-1) are engine operating parameters (23).

Total engine depot level overhaul cost includes both engine accessory overhaul cost and core-engine overhaul cost (23:4-6). However, since the decision to overhaul an engine accessory is made independently of a decision to overhaul a core-engine, only the overhaul costs for jet core-engines were considered in this study.

Research Objective

The objective of this research effort was to develop depot level cost estimating relationships for USAF jet core-engines utilizing both engine operating parameters and physical engine characteristics.

Research Hypothesis

It was hypothesized that cost estimating relationships may be developed to predict USAF core-engine depot level overhaul cost utilizing both engine operating parameters and physical engine characteristics.

CHAPTER II

RESEARCH METHODOLOGY

The research methodology employed in this study is designed to develop an acceptable cost estimating relationship (CER) for use as a predictor of depot level overhaul costs for military jet core-engines. The "core-engine," i.e., the basic engine without engine accessories, represents the primary area of concern.

DATA COLLECTION

Variables to Be Considered

The dependent variable (Y) used in this study denotes the depot level core-engine overhaul cost for military jet engines; hereafter referred to as overhaul cost. [1] The independent variables ($\bar{X}_i$'s) utilized are engine performance parameters and engine physical characteristics which were considered potential engine overhaul cost-drivers. Each of the independent variables is identified in Appendix A-2.

[1] The dependent variable utilized throughout this study was expressed as a core-engine overhaul cost/flight hour. "COST" may be freely substituted throughout this study in place of the term "overhaul cost."

These were considered potential cost-drivers based on a CER development study (3), an interview with Air Force Aero Propulsion Laboratory (AFAPL) engineers (16), and a pilot study conducted by the researchers concerning CERs (4).

The Air Force CER development study to determine, among other objectives, the best method of cost estimating jet engine production costs concluded that the use of jet engine physical characteristics (e.g., rotors, stators, blades, etc.) appear to be the best independent variables for consideration in CERs to estimate production costs. The problem with developing cost estimates for engine production ". . . stems [basically] from the fact that all engines are not simply scale ups or scale downs of a single basic design . . . [3:22]." The more imaginative the designers become, the more difference there will be from one engine to the next. However, even though engines may vary significantly in materials and technology, at the component level, all jet turbine engines have rotors, blades, and stators; otherwise they would be something else, e.g., ramjets or turborockets (3:21-24).

As previously discussed in Chapter I, present CER models for predicting USAF jet engine overhaul cost utilize exclusively, except for the one physical characteristic, engine dry weight, independent variables which are engine performance parameters. Based upon a review of the USAF/Air Force Logistics Command's (AFLC)

CERs for jet engine overhaul, the researchers felt that engine physical characteristics should also be included as independent variables (23). All jet turbine engines have certain basic physical characteristics in common which include:

1. A compressor assembly comprised of rotors, stators, and blades.

2. A turbine assembly comprised of rotors, stators, and blades.

3. A particular design of the combustion section, i.e., annular, semi-annular, can-annular.

Data Sample

Cost Data. Cost data for this study were provided by the AFLC's ICAD. The dependent variable referred to as COST represents the average jet core-engine depot level overhaul cost per flight hour. These data were accumulated for the engines listed in Appendix A-3 and represent fiscal years 1974, 1975 and 1976. Cost data were converted to equivalent 1976 dollars using price indices for engine costs from Air Force Regulation (AFR) 173-10.

Engine Sample. The engine sample consisted of the USAF type/model/series jet engines contained in Appendix A-3. As discussed in the Research Objective and Hypothesis section of Chapter I, present USAF jet engine CERs published by the AFLC considered primarily engine

operating parameters as the independent variables. Since the research hypothesis of this study was that CERs which considered both engine operating parameters and engine physical characteristics would produce better cost models, the engine sample utilized by the AFLC's ICAD to develop models published in AFLC Pamphlet 173-4 was adopted for comparative purposes.

Summary depot level core-engine overhaul COST was obtained from the AFLC ICAD by type/model/series engine and was used in a pilot study conducted by the researchers (4). The operating parameters input as independent variables in the research pilot study were incorporated in this study and the number of physical characteristics used as independent variables for each engine was increased. The increase resulted from discussions with AFAPL personnel (1). These additional characteristics were obtained from the illustrated parts breakdown manual and overhaul manuals for each specific engine. Physical engine characteristics which could not be located in either of the above sources were obtained from the Air Force Aero Propulsion Laboratory (AFAPL).

Within the engine sample, the data were grouped so that 80 percent of the engine data within each category was utilized to develop CER models and the remaining 20 percent of the engine data was withheld and used strictly for testing the predictive capability[2] of the

---

[2]A discussion of predictive capability will be presented later in the discussion.

developed CER models (13:38).

MODEL JUSTIFICATION

The objective was to develop CERs which were as uncomplicated as possible so that a cost estimate could be derived with minimum effort and time and at a minimum cost to the organization. Results of a pilot study conducted indicated that combinations of engine operating parameters and engine physical characteristics could be used to develop such CERs and that the methodology proposed herein would be most appropriate (4).

MODEL DEVELOPMENT

The CER models developed were relationships between one or more jet engine operating parameters and/or jet engine physical characteristics and overhaul costs. To find an acceptable CER, various CER models were developed using regression analysis techniques.

Regression Analysis

Regression analysis has the capability to identify potential relationships between two or more variables. Since the CERs developed are a relationship between two or more variables, regression analysis was assumed to be an appropriate modeling technique for this research, i.e., that the regression model would "fit" the

data (33:Ch. 13). The pilot study performed by the researchers supports this assumption (4). The assumption that historical data is indicative of and can be utilized to predict future occurrences is critical when using regression analysis as a technique for developing a predictive model.

Regression analysis is applicable to any finite number of independent variables (X_i's) if those variables can be assumed to have the qualities necessary to apply regression analysis. Those qualities are listed as assumptions of regression analysis in Appendix B.

Simple linear regression (SLR) involves the regression of a dependent variable (Y) on a single independent variable (X). Multiple linear regression (MLR) is an extension of SLR where Y is regressed on two or more X_i's. The forward stepwise inclusion MLR of the Statistical Package for the Social Sciences (SPSS) was used for CER model development (15:345-347). The SPSS forward inclusion MLR allows the user to specify the values of the three criteria just mentioned or to allow the package's built-in default criteria of F = .01, tolerance = .001, and maximum number of variables = 80 to be operative. In this study, the SPSS default criteria were employed. All variables not considered for inclusion in the regression model by the SPSS forward inclusion method were considered qualitatively by the researchers with respect to size of the sample and to the possible

effects of multicollinearity.[3] For a further discussion of regression analysis refer to Appendix B.

Forward inclusion is a method used in MLR for developing a regression equation. The independent variables included in the equation are included in the order of the highest contribution to explaining the variation of the regression model. The process continues until one or more of the following criteria are no longer met:

1. none of the F values (SPSS default criterion) of the remaining variables exceeds the minimum acceptable F value,
2. the tolerance criterion is not met,
3. the limit on the maximum number of variables to be included is exceeded.

Correlation Analysis

The dependent variable (Y) was regressed on various combinations of the independent variables (X_i's). The determination of which independent variables were considered for inclusion in the regression model was based on an analysis of the correlation coefficient (r) between each possible pair of independent variables. Only

---

[3]Multicollinearity is the situation in which there is a very high degree of correlation (approximately 0.90 or higher) between independent variables. As the correlation between independent variables approaches unity (representing a perfect linear relationship), the regression coefficients become less valuable for predictive purposes (5:616).

one of any pair of independent variables was considered if the r value was greater than .80 (14:407). The inclusion of succeeding independent variables into the CERs being developed was influenced by analysis of the values of first order partial correlations.

Correlation analysis is designed to determine the extent of the linear relationship between variables. Correlation analysis can point out spurious[4] as well as valid useful relationships, however, a correlation from chance causes may be found between variables which may have no logical relationship to each other. The coefficient of correlation (r) will take on a value between -1 and +1. A value of +1 indicates a perfect linear positive or proportional relationship between two variables. A value of -1 also indicates a perfect linear relationship between two variables; however, the relationship is negative or inverse. A correlation value of zero indicates no linear relationship between the variables correlated. It must be noted, however, that correlation does not necessarily imply that a cause-and-effect relationship exists (5:542-558; 15:276-280; 10; 34).

Correlation between any two variables is called bivariate, simple correlation, or zeroth order partial correlation. The coefficient of correlation of a bivariate correlation is represented herein as r_o.

---

[4]A spurious relationship between two variables is indicated when the value of their first order partial correlation coefficient is some measure lower than the value of their bivariate correlation coefficient (9:317).

First order partial correlation measures the relative importance of each of the independent variables separately in explaining variations in the dependent variable while at the same time controlling for the variations in another independent variable. Computations of first order partial correlation (represented herein as r_1) may be performed to measure the extent to which the variation in the dependent variable, which was not explained by all the other independent variables, is explained by the addition of the new independent variable (9:317; 34).

<u>Correlation Evaluation Rule 1.</u> Bivariate and first order partial correlation analysis was performed on each of the engine categories listed in Appendix A-4. To aid the analysis of bivariate correlation coefficients, a bivariate coefficient of correlation value of $r_o \geq \pm .7000$ was arbitrarily established as an initial criterion to select those independent variables which appeared to have the strongest linear correlation with COST. This criterion is herein referred to as Rule 1.

<u>Correlation Evaluation Rule 2.</u> A consideration in choosing variables for developing a CER is to exclude one of a pair of independent variables which exhibits a high r_o between them. This action is taken to avoid the problems of multicollinearity. It was arbitrarily determined that if a bivariate correlation coefficient between any two independent variables was equal to or greater than $\pm .8000$, the use

of these two variables together in an MLR model would introduce the effects of multicollinearity. To aid in the selection of variables to be used in an MLR, the researchers defined the coefficient of correlation ranges between paired independent variables as follows:

Range of r_o	Meaning	Action
0 to $\pm$ .50	Little or no linear correlation between the paired variables.	Regress both variables with COST.
$\pm$ .50 to $\pm$ .79	Moderate linear correlation between the variables.	Regress one of the variables with COST. If the number of variables available for regression is small, an arbitrary judgment as to the inclusion of the second variable is appropriate.
$\pm$ .80 to $\pm$ 1.00	High linear correlation between the variables.	Regress only one of the pair with COST.

This bivariate coefficient of correlation evaluation criterion is referred to as Rule 2 hereafter.

Correlation Evaluation Rule 3. To aid the analysis of first order correlations, searching for spurious bivariate correlations, the researchers defined the following heuristic rule hereafter referred to as Rule 3:

1. If the absolute value of r_1 is less than the absolute value of r_o minus .2000, the bivariate correlation value would be

considered to represent a spurious relationship due to the effects of the r_1 control variable.

2. If the algebraic sign of the r_o value was opposite the sign of the r_1 value, the r_o value would be considered to represent a spurious relationship due to the effects of the r_1 control variable.

The application of heuristic Rules 1, 2, and 3 aided the researchers in the analysis of bivariate correlations between the dependent variable COST and the independent variables listed in Appendix A-2, the analysis of bivariate correlation between pairs of independent variables, and the first order correlation of COST with each independent variable.

Model Categorization

The CERs which have been developed by the Independent Cost Analysis Division (ICAD) of the AFLC categorized jet engines as follows (23):

1. TURBOJETS
2. TURBOFANS

However, as a result of the pilot study conducted, it was concluded that there may be other categorizations of jet engines which might produce statistically significant models and provide a more accurate estimate of jet engine overhaul cost. The additional categories identified were:

1. ALL ENGINES
2. TURBOJETS (with afterburner)
3. TURBOJETS (without afterburner)
4. TURBOFANS (with afterburner)
5. TURBOFANS (without afterburner)

MODEL VALIDATION

Regression models constructed were subjected to standard statistical tests to determine the statistical significance and acceptability of the model. F-tests were performed on the overall regression model and on the net contribution of the individual independent variables (X_i's) to the explained variation. The X_i's were tested simultaneously using the Bonferroni technique to compute equivalent-alpha (α) values (14:257). The coefficient of determination (R^2) and the standard error of the estimate were also considered for model acceptability. Confidence intervals were constructed for each model at the 95 percent confidence level to test the predictive capability of the model and as an indicator of its utility. If the overall model was found to be significant at the 95 percent confidence level, and the coefficient of determination (R^2) was greater than .8000, an arbitrary choice (14:92), the model was considered acceptable. Acceptance criteria for the standard error of the estimate and the accuracy and reliability of the model's predictive ability were not established since

neither the ICAD nor the AFAPL have established a minimum acceptance level for these criteria (1; 18; 32).

Predictive Capability

Cost model predictability was examined for each CER developed. Once a model had been developed using 80 percent of the sample data and met the R^2 and confidence level criteria previously discussed, an average percent predictability was computed. Using a test engine, the average percent predictability was computed by dividing the absolute numerical difference between the test engine's actual COST and predicted COST by the actual COST:

$$\frac{\text{actual COST - predicted COST}}{\text{actual COST}} \times 100$$

The maximum acceptable predictability level was established by the researchers as 50.0 percent.

SUPPORT OF RESEARCH HYPOTHESIS

If the CER model(s) met the criteria for statistical significance and acceptability presented in the Model Validation section, the research hypothesis was considered supported and the CERs developed were considered acceptable.

SUMMARY LIST OF ASSUMPTIONS

1. The independent variables chosen are cost-drivers with respect to overhaul cost and that they are logically supportable.

2. The relation between the dependent variable and independent variables is linear.

3. Regression analysis is an appropriate modeling technique.

4. Historical data is indicative of and can be utilized to predict future occurrences.

5. The independent variables satisfy the assumptions of regression analysis.

SUMMARY LIST OF LIMITATIONS

1. CERs generally cannot be used to estimate costs of radically different new systems since CERs are developed from historical experience (11:7).

2. The accuracy of the cost data is limited by the accuracy of the reporting system.

3. The researchers were uncertain as to whether the independent variables identified and used to develop jet core-engine overhaul CERs represented a complete list of all possible cost drivers.

4. Complete data records were not available for all engines.

5. The independent variables chosen were not totally applicable to each engine category identified for investigation.

CHAPTER III

DATA ANALYSIS AND FINDINGS

INTRODUCTION

Chapter III presents the analysis and findings of the research, which was conducted to develop parametric cost estimating relationships[1] (CERs), for predicting core-engine depot level overhaul costs, for USAF jet engines. Initial discussion will focus on the independent variables utilized in the study and the analysis and findings of a bivariate and first order partial correlation analysis which was performed.

Following the presentation of the findings of the correlation analysis, a two-part discussion of the analysis and findings of the CERs which were developed is provided. The first part of the discussion, pertaining to CER development will focus on the models developed, an analysis of the coefficient of determination (R^2), confidence level, and percent predictability obtained for each engine category model. The second part will cover an analysis of the distribution of R^2 values for models developed in each engine category and an analysis of cost models developed by use of a modified multiple

[1]Throughout this chapter, the terms "cost model" or "model" should be interpreted to mean cost estimating relationship.

linear regression (MLR)[2] program. The emphasis of this discussion is on the expected R^2 value, the average percent predictability, and average percent efficiency computed for cost models derived for each of the engine categories: ALL ENGINES, TURBOJETS, and TURBOFANS.

INDEPENDENT VARIABLE LIMITATIONS

The independent variables identified for use in this study are identified in Appendix A-2. The assumption made was that the independent variables chosen were cost-drivers of core-engine overhaul cost and that their utilization in a CER to predict this cost could be logically supported. However, limitations regarding the independent variables were uncovered once the data collection effort was complete. First, it was noted that complete data records were not available for all engines. The two independent variables that contributed to the lack of complete data records were variable X14 (NORMTIT) and variable X15 (MILTIT). Complete data elements for these variables were not obtained because, depending on the specific engine, either NORMTIT or MILTIT ratings were not published. Because of the inability of the researchers to collect data on these two independent variables for all engines, these variables were not considered in the development of any CER.

[2]The modified MLR program is discussed in Appendix D.

The second limitation noted, concerning the independent variables, related to whether a specific independent variable should be considered in the development of a specific CER. It was found that the independent variable X13 (MAXT) had been obtained only for those engines which are equipped with afterburners. Therefore, variable X13 was considered only in model categories (1) TURBOFANS (with afterburner) and (2) TURBOJETS (with afterburner). Variable X19 (MAXSFC) followed the same pattern as variable X13, i.e., a data element was obtained for variable X19 for only those engines equipped with an afterburner. Consequently, variable X19 was only considered for models developed for TURBOFANS (with afterburner) and TURBOJETS (with afterburner). Variable X22 (FBR) is unique to turbofan engines. Consequently, it was considered in CERs for TURBOFANS, TURBOFANS (with afterburner), and TURBOFANS (without afterburner).

Appendix C-1 lists the independent variables considered in CER development by engine category. Prior to the initial development of CERs using the MLR technique, the independent variables to be considered in each engine category were subjected to bivariate and first order partial correlation analysis.

CORRELATION ANALYSIS

Bivariate correlation (r_o) analysis between the dependent variable COST and the independent variables, and between each pair

of independent variables was performed by engine category. Additionally, first order partial correlation (r_1) analysis was performed between the dependent variable COST and each independent variable that met the Rule 1 criterion[3] defined in Chapter II. However, due to the small number of engines included in four of the seven engine categories as shown in Appendix A-4, correlation analysis was limited to the categories ALL ENGINES, TURBOJETS, and TURBOFANS.

ALL ENGINES

The engines included in the correlation analysis for the category ALL ENGINES are listed in Appendix C-2.1. Of the 22 independent variables considered (Appendix C-1) in this category as possible core-engine cost drivers, no independent variable displayed a bivariate coefficient of correlation with the dependent variable COST equal to or greater than $\pm .7000$. Therefore, the minimum level, r_o, for Rule 1, was arbitrarily lowered first to $\pm .6000$ and then to $\pm .5000$. At the $r_o \geq \pm .6000$ level, two of the 22 independent variables, X6 (L) and X25 (AFHBO), displayed r_o values that met or exceeded the minimum requirement. At the $r_o \geq \pm .5000$ level, three additional variables, X8 (LD), X16 (MAXTIT), and X23 (EPR), were

---

[3]Extensive reference is made throughout the Correlation Analysis section of this chapter to "Rules 1, 2, and 3." These "Rules" were introduced in the Correlation Analysis section of Chapter II to assist the researchers in evaluating the results of bivariate and first order partial correlation analysis.

identified as possibly having a meaningful linear relationship with COST. Appendix C-3 lists the five independent variables that displayed a bivariate coefficient of correlation with COST equal to or greater than $\pm$ .5000.

Of the five variables which met the modified Rule 1 minimum of $\pm$ .5000, variable X6 exhibited the highest direct correlation with COST (.6721) and the variable X25 displayed an inverse r_o relationship with COST of -.6278. The remaining three variables (X8, X16, and X23) displayed r_o values of .5235, .5484, and .5847 respectively. Appendix C-4 provides a tabular summary of the bivariate coefficients of correlation between the five variables mentioned above which met the modified Rule 1 r_o criterion.

Rule 2 was used to classify the strength of the linear relationship between independent variables. Of the 10 pairs of independent variables analyzed, one pair, X16 (MAXTIT) with X23 (EPR), was classified as having a high linear relationship. Five pairs of independent variables, X8 (LD) with X6 (L), X6 with X16, X6 with X23, X6 with X25 (AFHBO), and X16 with X25, were classified as having a moderate linear relationship, and four pairs, X8 with X-16, X8 with X23, X8 with X25, and X23 with X25, were classified as having no meaningful linear relationship between the paired variables (refer to Appendix C-4).

A first order partial correlation analysis was performed with each bivariate combination of the dependent variable, COST, and the five independent variables which met the modified Rule 1 criterion. Appendix C-5 presents the results of that r_1 correlation. The researchers were interested in determining those r_0 correlations with the dependent variable which could be shown by first order partial correlation analysis to be spurious relationships. To accomplish this analysis, the evaluation criteria of Rule 3, as defined in Chapter II, was applied.

X6 (L)

The first order partial correlation of X6 with COST, controlling for each other independent variable, is shown in Appendix C-5. The r_0 value of X6 with COST was computed to be .6721. Appendix C-5 indicates that this bivariate correlation is upheld by first order partial correlation when controlling for all other variables except X23 (EPR) and X25 (AFHBO). Both X23 and X25 are among those independent variables initially selected as having the best linear correlation with COST ($r_0 \geq \pm .5000$).

X8 (LD)

The bivariate correlation coefficient of X8 and COST (.5235) was found to be supported by all control variables except X6 (L) and X25 (AFHBO). The variables X6 and X25 were among those variables

which met the modified Rule 1 criterion.

X16 (MAXTIT)

The first order partial correlation of X16 with COST, controlling for the effects of all other independent variables, indicated that the r_0 value of .5484 was supported by all variables except X6 (L), X23 (EPR), and X25 (AFHBO). Variables X6, X23, and X25 are among those independent variables that were considered to have the best linear correlation with COST.

X23 (EPR)

Analysis of the first order partial correlation of X23 with COST, while controlling for all other variables, depicts the same pattern evidenced by the r_1 values of X6 (L), X8 (LD), and X16 (MAXTIT). That is, the first order partial correlation coefficient of X23 with COST was supported by all control variables except those control variables which met the modified Rule 1 criterion.

X25 (AFHBO)

The pattern evidenced throughout the analysis of first order partial correlations of Rule 1 variables has indicated that variables that displayed the best bivariate correlation with COST do not support bivariate correlations of other variables with COST when used as control variables. The first order partial value of X25 (-.6278)

is upheld when controlling for all other variables except X6 (L) and X16 (MAXTIT).

Summary of ALL ENGINES Correlation Analysis

The results of the first order partial correlation analysis of Rule 1 variables suggests that the bivariate correlation of each independent variable with COST is a valid correlation. Application of Rule 2 in selecting variables to regress in an MLR suggested which combinations of independent variables should not be regressed together if the effects of multicollinearity on the regression equation are to be avoided. Appendix C-6 presents representative combinations of Rule 1 independent variables that may produce the simplest CER while avoiding the affects of multicollinearity between the independent variables.

TURBOJETS

Engines included in the correlation analysis performed on the engine category TURBOJETS are listed in Appendix C-2.2. Of the 22 independent variables considered (Appendix C-1),it was found that, similar to the bivariate correlation analysis performed on the category ALL ENGINES, no independent variable displayed a bivariate coefficient of correlation with COST greater than the Rule 1 minimum of $\pm .7000$. The researchers therefore chose to arbitrarily lower the Rule 1 criterion for r_o from $r_o \geq \pm .7000$ to $r_o \geq \pm .5000$. By

lowering the r_o value, seven of the 22 variables correlated with COST displayed a coefficient of correlation whose value exceeded the lower r_o limit. These variables are listed in Appendix C-7. Appendix C-8 provides a summary of the bivariate coefficients of correlation between each pair of Rule 1 variables.

Rule 2 was applied to classify the degree of each linear relationship. Of the 21 pairs of independent variables correlated, 12 pairs were classified as displaying a high linear correlation: X1 (CR) with X2 (CB), X1 with X3 (TR), X1 with X8 (LD), X1 with X23 (EPR), X2 with X3, X2 with X23, X3 with X6 (L), X3 with X8, X3 with X23, X6 with X8, X6 with X23, and X8 with X23.

First order partial correlation was performed on each bivariate combination of COST with the seven variables which met the modified Rule 1 criterion. Appendix C-9 displays the result of that correlation. Again, the objective was to determine whether the r_o correlations depicted in Appendix C-7 were indicative of true linear correlations or whether they were the results of spurious relationships caused by the action of the control variable. Rule 2 was applied to aid the researchers in this analysis.

<u>X1 (CR), X2 (CB), X3 (TR), X6 (L), X8 (LD), and X23 (EPR)</u>

First order partial correlations analysis of the listed variables suggested that the r_o value for each of these variables with

COST may be spurious. In each case, the control variables, which did not support the r_o correlation for the listed variables with COST, were primarily those variables whose bivariate correlation with COST was equal to or greater than $\pm .5000$.

X25 (AFHBO)

The first order partial analysis of variable X25 (AFHBO) suggests that the r_o value is a true indication of its correlation with COST. Only two variables did not uphold the r_o value of X25. They were variables X6 (L) and X8 (LD); both of which displayed r_o values equal to or greater than $\pm .5000$.

Summary of TURBOJET Correlation Analysis

The results of the bivariate correlation analysis of the independent variables with COST found that none of the independent variables displayed an r_o value equal to or greater than $\pm .7000$. This fact caused the researchers to arbitrarily lower the Rule 1 minimum r_o value to $\pm .5000$. The lowering of the r_o value allowed seven variables to meet the modified Rule 1 criterion for further first order partial correlation analysis. Of the seven variables analyzed, first order partial correlation analysis suggested that only one variable, X25 (AFHBO), exhibited a nonspurious bivariate coefficient of correlation with COST. In each case the control variables which did not uphold the r_o correlation were those variables which also met the

modified Rule 1 criterion. The one exception was variable X9 (W). When X9 functioned as the control variable for variables X1 (CR) and X3 (TR), it did not support either the X1 or the X3 r_o values with COST.

TURBOFANS

The engines included in the correlation analysis for the category TURBOFANS are listed in Appendix C-2.3. Of the 23 independent variables considered (Appendix C-1), 14 variables were found to display a bivariate coefficient of correlation with COST above $\pm$ .7000. The independent variables which met the Rule 1 criterion and which would be subjected to further correlation analysis are listed in Appendix C-10.

A bivariate correlation was performed between each combination of independent variables whose bivariate correlation with COST met the Rule 1 criterion (refer to Appendix C-11). Rule 2 was applied to assist in the analysis. Of the 123 pairs of variables examined, 116 pairs were classified as having high linear relationships, 7 pairs were classified as having moderate linear relationships, and none of the variable pairs examined had an r_o less than $\pm$ .6000. Additionally, one pair of the independent variables examined, X6 (L) with X8 (LD), exhibited a perfect linear relationship (r_o = 1.0000). Appendix C-12 displays the results of the first order partial correlation performed between COST and each Rule 1 variable.

Rule 3 was applied to assist the researchers in determining which r_o correlations with COST were true indications of linear relationships and not spurious relationships influenced by the action of the control variable. The first order partial correlation analysis indicated that the bivariate correlation of variable X4 (TB) with COST was the only r_o generally supported. Two control variables, however, did not support X4's r_o with COST: X10 (TW) and X23 (EPR). Like the results of the correlation analysis performed on other engine categories, the variables which did not support the r_o value of X4 were also Rule 1 variables. All other independent variables, whose r_o with COST was tested by the application of Rule 3, were considered to display spurious relationships with COST.

Summary of TURBOFANS Correlation Analysis

The TURBOFANS category had 10 variables which met the Rule 1 criterion and 116 of 123 pairs of Rule 1 variables that were considered to have high linear relationships. However, although the variables of this category displayed strong results after the a[illegible]cation of Rules 1 and 2, only one variable, X4 (TB), withstood a first order partial correlation challenge of its bivariate correlation value with COST.

Findings of Bivariate and First Order Correlation Analysis

Bivariate and first order partial correlation analysis of the variables in the engine categories ALL ENGINES, TURBOJETS, and

TURBOFANS were performed to assess the presence of the linear relationships between the category variables. The Rule 1 criterion defined in Chapter II established a measure to classify or select those variables which would be considered to have the most meaningful linear relationship with COST. Of the three engine categories, only the category TURBOFANS produced any bivariate coefficients of correlation of the independent variables with COST which met the Rule 1 criterion ($r_o \geq \pm .7000$). This prompted the researchers to arbitrarily reduce the Rule 1 minimum r_o value from $\pm .7000$ to $\pm .5000$ for bivariate correlation analysis of the engine categories ALL ENGINES and TURBOJETS.

Although the Rule 1 minimum was modified downward for the engine categories ALL ENGINES and TURBOJETS, the first order partial correlations for these categories supported the Rule 1 bivariate correlations with COST. Conversely, the application of Rule 3 criteria to the first order partial correlation values for the category TURBOFANS did not support the bivariate correlations with COST. The bivariate correlations had indicated that there were strong linear relationships between COST and the independent variables of this engine category.

The results of the preceeding correlation analysis were considered in the second portion of the cost model development: variable selection for CER development.

INITIAL COST MODEL ANALYSIS

As described in Chapter II, the multiple linear regression technique was utilized to develop cost estimating relationships for the engine categories listed in Appendix A-4. The CER selected from each category was chosen based on the highest coefficient of determination ($R^2 \geq .8000$) at an overall model confidence level of 95 percent. Once the model was chosen for each of the engine categories, the average percent predictability for each model was calculated utilizing the test engines listed in Appendices C-2.1 through C-2.7. The arbitrary maximum value of percent predictability a model could display, and be considered acceptable, was 50.0 percent.

ALL ENGINES

The CER for the engine category ALL ENGINES was developed utilizing the data for engines listed in Appendix C-2.1 and the independent variables indicated in Appendix C-1 and had an $R^2 = .9003$. Appendix C-13 lists the independent variables and their coefficients of the cost model which met both the R^2 and the confidence level criteria. Even though the model chosen was significant at the 95 percent confidence level, only one of the 14 independent variables, X16 (MAXTIT), made a significant marginal contribution

to the model's overall explanatory power.[4]

Appendix C-14 displays the calculation of the model's predictive capability using the test engines. From Appendix C-14 it may be observed that the model chosen which had the highest R^2 at a 95 percent confidence level did not produce an average percent predictability within the acceptable (0 to 50.0 percent) range.

TURBOJETS

Application of the MLR technique using data for engines listed in Appendix C-2.2 did not produce a CER which met both the R^2 and significance level criteria required for initial model selection. The R^2 obtained for a TURBOJET model was .6101 (refer to Appendix C-15).

TURBOFANS

Appendix C-16 displays the CER developed for the engine category TURBOFANS with an R^2 of 1.0000. Each of the five coefficients of the independent variables was found to contribute significantly to the model's overall explanatory power.

Appendix C-17 displays the predictive capability of the model, which was calculated by using data for the test engine. The

---

[4]Of the 14 variables in the model only X16 appears to be a significant driver of cost, i.e., X16's marginal contribution to the R^2 value of .9003 was the only contribution to exhibit statistical significance when subjected to an F-test.

overall average percent predictability of the model was calculated to be 220.0 percent which exceeded the 50.0 percent maximum established for model acceptability. Of the two test engines utilized, one engine (TF30-P-7) produced a percent predictability of 41.0 while the other engine (TF39-GE-1) produced a percent predictability of 399.0.

TURBOFANS (with Afterburner)

Data for engines shown in Appendix C-2.4 were utilized to develop the CER listed in Appendix C-18. The model has one independent variable and displays an $R^2 = .9976$. The predictive capability of this model was calculated to be 72.0 percent (refer to Appendix C-19).

TURBOFANS (without Afterburner)

Regression of data for the engines listed in Appendix C-2.5 produced the CER shown in Appendix C-20. At the 95 percent confidence level, this model displayed an $R^2 = .9999$. Additionally, the coefficients of the two independent variables included in the model were found to contribute significantly to the model's overall explanatory power.

Appendix C-21 displays the percent predictability calculated for the overall model. The single test engine produced a 62.35 percent predictability.

TURBOJET (with Afterburner)

The CER listed in Appendix C-22 was developed using the data for engines shown in Appendix C-2.7. The model displays an R^2 value of .9869; however, only one of the coefficients of the four independent variables was found to make a significant marginal contribution to the model's overall explanatory power. The overall average predictive capability of this model was calculated to be 285.0 percent (Appendix C-23). However, of the two test engines utilized, one engine (J57-p-21) produced a predictability of 25.0 percent and the other engine (J85-GE-13) produced a predictability of 545.0 percent.

Findings

Cost estimating relationships were developed for each of the seven engine categories using the MLR procedure. However, of the seven models developed, only five met the minimum initial acceptance criteria of $r^2 \geq .8000$ at the 95 percent confidence level. Of the five models that met these criteria, no model displayed an overall average percent predictability equal to or less than the 50 percent maximum set by the researchers. The utility of a model as a predictor of cost is a subjective determination by the user of a model based on the accuracy of a prediction he is willing to accept. However, the accuracy and usefulness of a model as a predictor is inversely related to the absolute value of the model's percent, or average percent, predictability. Appendix C-24 provides a summary

of these statistics.

Although none of the cost models met the percent predictability criterion, two models were close. They were TURBOFANS (with afterburners) and TURBOFANS (without afterburners) which displayed average overall model predictabilities of 72.0 percent and 62.35 percent respectively.

The engine category TURBOFANS displayed a percent predictability of 220.0. However, one of the two test engines (TF30-P-7) displayed a 41 percent predictability while the second test engine (TF39-GE-1) displayed a 399.0 percent predictability.

Of the three major engine categories (ALL ENGINES, TURBOJETS, and TURBOFANS), only two, ALL ENGINES and TURBOFANS, produced cost models which met both the R^2 and confidence level criteria. Of the three TURBOJET subcategories (TURBOJETS, TURBOJETS (with afterburner), and TURBOJETS (without afterburner)) only the category TURBOJETS (with afterburner) produced a CER that met both the R^2 and confidence level criteria.

All CERs developed that met the R^2 and confidence level acceptance criteria were analyzed with regard to the makeup of the independent variables that comprised each model; numbers of engine operating parameters vs. numbers of physical engine characteristics. Of the 14 independent variables that make up the ALL ENGINES cost model, 10 of 14 variables are physical engine characteristics. The

CER, shown in Appendix C-16, for the TURBOFANS category is comprised of five independent variables of which two, X18 (*MILSPC*) and X27 (AFH), are engine operating parameters, two, X2 (CB) and X4 (TB), are engine physical characteristics, and the fifth variable, X10 (TW) is a combination of both variable categories.

The remaining three engine CERs, TURBOFANS (with afterburner), TURBOFANS (without afterburner), and TURBOJETS (without afterburner) each had equal shares of both operating parameters and physical characteristic variables. The exception was the TURBOFAN (with afterburner) model which had only a single independent variable, X23 (EPR).

The indication at this stage of the analysis is that physical engine characteristics contribute as well to a model's overall explanatory power as do engine operating parameters. This indication is based on a simple ratio of the total number of independent variables in the five models discussed. Sixty percent of the independent variables in the developed cost models are physical engine characteristics (Appendices C-12, C-14, C-15, C-19, C-21).

COST MODEL ANALYSIS

The initial models developed for the ALL ENGINES and TURBOFANS categories had R^2 values of .9003 and 1.0000 respectively. However, the R^2 value for the model developed for

TURBOJETS was only .6101.

The marginal contribution of the variables to the model in the ALL ENGINES and TURBOJET categories were insignificant except for the variable X16 (MAXTIT) in the ALL ENGINES model.

The high R^2 values for the ALL ENGINES and TURBOFANS models, the low R^2 value for the TURBOJETS model, small sample sizes for all categories, and the insignificant marginal contributions of the variables suggested that the data may be of questionable reliability, at least in the ALL ENGINES and TURBOJETS categories. It also appeared that the validity of the models developed might be very sensitive to the particular engine data cases selected for use in building the model and those used in testing the model. It was questionable whether the data samples were representative of the population. Hence, a more complete analysis was undertaken to determine the validity of the data samples with respect to the quality of a model developed using the data. The extended analysis was performed on the ALL ENGINES, TURBOJETS and TURBOFAN categories.

Extended Cost Model Analysis

Method of Analysis. To determine whether the data in each category could be expected to yield a model with an $R^2 \geq .8000$, a series of multiple linear regression (MLR) models were developed for each data category. In each category, an engine record was removed

from the data file and an MLR model developed using the remaining cases. The procedure was followed until each engine case had been sequentially removed and a model developed. The distribution of R^2 values for the models in each category were studied with respect to the mean (expected) value, the range, and the number of models with an $R^2 \geq .8000$. If the expected R^2 value was equal to or greater than .8000, the range of R^2 values was small, and the number of R^2 values less than .8000 was small, the data could be considered representative of the population. Further, an R^2 value equal to or greater than .8000 for a model developed from such data could be considered reliable. Further, that when subjected to an F-test, the lack of statistical significance of a model, with an $R^2 \geq .8000$, could be attributed to the lack of statistical leverage resulting from the large number of independent variables in the model and the small data sample size (10). Then variables were chosen to construct a representative model of the data by comparing variables in models with R^2 values closest to the expected R^2 value of each respective data category. The number of variables chosen would be equal to or less than the number appearing in the models compared, i.e., if there were 17 variables in each of the models compared, then 17 would be chosen for use in the remainder of the analysis. A smaller number of variables were chosen if the data sample was not large enough to ensure adequate degrees of freedom in the denominator when computing a model

F-value with all variables in the model.

A computer program (MREG), explained in Appendix D, was used to randomly divide the data into a percentage for model development and a percentage for testing the developed model with respect to expected predictive efficiency and its expected predictive capability. A predictive efficiency of 95 percent and a predictive capability of from 0 to 50 percent was desirable. In order to attain a representative distribution, MREG was run and models built and tested 1000 times for each of the data categories.

The number of variables used was then reduced so that the value of n-p (refer to Appendix B, F-tests) could be increased resulting in greater statistical leverage in the model and causing the confidence intervals (refer to Appendix B; confidence intervals) for the test cases to be narrowed.

MREG was run another 1000 times with a reduced number of variables for each data category. This second run was to determine whether or not the model's expected predictive efficiency would remain at a level equal to or greater than the level attained with a greater number of variables and wider confidence intervals. If the original level and the subsequent level were both greater than 95 percent, it could be considered that a model developed from the particular data category using all the data would remain at least 95 percent efficient (10).

The variables selected for elimination in the second iterations were chosen by comparing the variables in the same models used to initially choose the variables for further analysis and by the results of previous correlation analysis. The point at which the R^2 value in each model exceeded .8000, the minimum acceptable R^2 was considered a cutoff point, and any variables which entered the model beyond that point were eliminated.

<u>ALL ENGINES</u>. As shown in Appendices A-4 and C-2.1, the ALL ENGINES category data sample contained 28 engine cases. Twenty-eight separate models were developed utilizing the independent variables listed in Appendix C-25 with a different engine case eliminated from the data each time. The R^2 values ranged from .7928 to .9185. The R^2 of .7928 was the only value below .8000 and occurred when the engine data for the TF39-GE-1 engine was eliminated from the sample. The range of the remaining values was .8761 to .9185; a spread of .0424. The mean R^2 value for the distribution was .8901.

Five models whose R^2 values were clustered about the mean were used to choose the initial variables for analysis. These variables are listed in Appendix C-25. MREG was run 1000 times utilizing the variables chosen and, each time, 80 percent of the data sample was used to construct the models. It was determined from the models developed that an R^2 value of .9138 with a range of .3593,

a percent predictability of 417.69 percent with a range of 72,309, and a predictive efficiency of 96.56 percent with a range of 88.89 could be expected of a model developed from the data utilizing the variables referenced.

When the variables used were reduced to those listed in Appendix C-26, the degrees of freedom were increased from four to seven, the confidence intervals were reduced by 14 percent, and the expected predictive efficiency dropped to 94.03 percent, a decrease of 2.53 percent. Appendix C-27 provides comparative statistics of R^2, percent predictability, and percent efficiency at degrees of freedom equal to five and sixteen.

TURBOJETS. The TURBOJETS category data sample is comprised of 19 engine cases as indicated in Appendix C-2.2. Nineteen models considering the independent variables listed in Appendix C-28 were developed with R^2 values ranging from .7004 to .9996. Seven of the values ranged from .7004 to .7848, two of the values were between .8000 and .9000, and the remaining ten ranged from .9071 to .9996. The mean R^2 value was .8651.

The six models closest to the mean were compared to choose the initial variables in this category. Because of the smaller sample size n = 19, 90 percent of the sample was used for model development in each of the 1000 models. The models yielded an expected R^2 value of .6928 with a range of .8655, an expected percent predictability of

289 percent with a range of 67,608, and an expected predictive efficiency of 93.32 percent with a range of from zero to 100 percent.

When the variables were reduced to those listed in Appendix C-29, the degrees of freedom were increased from four to seven, the confidence intervals narrowed by 15 percent, and the expected predictive efficiency was reduced to 92.74 percent, a drop of .58 percent.

Appendix C-30 provides a comparison of the R^2, percent prediction, and percent efficiency for degrees of freedom four and seven respectively.

<u>TURBOFANS.</u> The nine turbofan engines are shown in Appendix C-2.3. The range of R^2 values for the nine models developed was from .9910 to 1.0000 with a mean value of .9988.

Since the range of R^2 values was so small (.009), all of the models were compared to choose the initial variables that are listed in Appendix C-31.

Again, because of the small sample size, 90 percent of the sample was used to build each of the 1000 models. It was determined that an R^2 value of .9986 with a range of .0027, a percent predictability of 54.25 percent with a range of 127, and a predictive efficiency of 87.70 percent with a range of from zero to 100 percent could be expected of a model developed from the variables and data sample used.

The number of variables were also reduced (Appendix C-32) in this category increasing the degrees of freedom from one to four and reducing the confidence intervals by 75 percent resulting in a diminished predictive efficiency of 86.82 percent, a reduction of .88 percent. Appendix C-33 provides a comparison of R^2, percent predictability, and percent efficiency for degrees of freedom one and four.

Final Models. An MLR model was then developed for each category using the representative variables initially chosen for further analysis, (Appendices C-25, C-26, and C-31). To develop these models, the SPSS forward inclusion technique was used and all of the engine data cases were used for each respective category. The model for each category is listed in Appendices C-34, C-35, and C-36.

Findings

Many points in the data analysis indicate that the turbofan data, when combined with the turbojet data, caused the high R^2 for the ALL ENGINES category.

The R^2 value was .9003 for ALL ENGINES, 1.0000 for TURBOFANS, and .6100 for TURBOJETS. Only one of the variables in the ALL ENGINES model made a significant marginal contribution while all of the variables in the TURBOFANS model made extremely significant contributions. None of the variables in the TURBOJETS

category made a significant contribution.

The distribution of R^2 values for ALL ENGINES had one value less than .8000, and the range of the remaining values was only .04. The values for TURBOFANS were tightly clustered with high R^2 values covering a range of only .0090. However, the R^2 values for TURBOJETS were not distributed as well; the range was .2992 with seven values below .8000, two values between .8000 and .9000, and the remaining 10 values were greater than .9000.

Further analysis of the 1000 models developed in each category supported the observation. The mean R^2 value was .9138 for ALL ENGINES, .6928 for TURBOJETS, and .9986 for TURBOFANS. The percent predictability was 417.69 with a range from 14.99 to 72324 for ALL ENGINES, 288.99 with a range of 2.18 to 67,608.20 for TURBOJETS, and 54.25 with a range of .93 to 127.66 for TURBOFANS.

One reason for the TURBOFAN data's influence on the explanatory power (R^2 value) of the ALL ENGINES model could be due to more accurate cost figures in the data sample. Another reason might be that the turbofan data is more representative of its population or a combination of both reasons may apply.

A comparison of the distributions of R^2 values for each category suggests that the data in the TURBOJETS category is not stable throughout the data sample while the data in the TURBOFANS

appears to be stable. Since the ALL ENGINES category is comprised of the turbojet and turbofan data, the relationships of turbofan data with COST may be masking and assisting a weak relationship between turbojet data and COST.

The final chapter will present the summary, conclusions and recommendations resulting from the research. In addition, recommendations for extended research are provided.

CHAPTER IV

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS FOR FUTURE STUDY

SUMMARY

The objective of this research effort was to develop depot level cost estimating relationships (CERs) for USAF jet core-engines utilizing both engine operating parameters and physical engine characteristics. Present CERs published by the Air Force Logistics Command's (AFLC) Independent Cost Analysis Division (ICAD) utilize primarily engine operating parameters as variables to predict overhaul cost (18).

The study's scope was limited to the development of CERs for jet core-engines. This limitation is in contrast to present AFLC CERs which attempt to predict total jet engine overhaul cost (i.e., the core-engine plus engine mounted accessories). Since engine accessory overhaul may be performed independent of core-engine overhaul, only CERs for jet core-engine depot level overhaul were investigated.

The engines utilized by the AFLC to develop CERs published in AFLC Pamphlet 173-4 were selected as the data base engines for

this research effort. Additionally, engine physical characteristics were considered as possible overhaul cost drivers.

Seven engine categories were initially selected for CER development. They were:

ALL ENGINES

TURBOJETS

TURBOFANS

TURBOFANS (with afterburner)

TURBOFANS (without afterburner)

TURBOJETS (with afterburner)

TURBOJETS (without afterburner)

Of the original seven engine categories investigated, only three categories (ALL ENGINES, TURBOJETS, and TURBOFANS) were selected for final model development. Final models were developed by utilizing the technique of multiple linear regression. The researchers established model acceptance criteria of $R^2 \geq .8000$, at a 95 percent confidence level, with a predictive efficiency of not greater than 50 percent, and an overall model efficiency not lower than 95 percent. Eighty percent of the ALL ENGINES data base records were randomly selected for model development, and 90 percent of the TURBOJETS and TURBOFANS data base records were randomly selected. The values 80 and 90 percent were decided upon by the researchers after considering the sample sizes, n, in each

category (n = 28 for ALL ENGINES, n = 19 for TURBOJETS and n = 9 for TURBOFANS). Models were developed in each category using the randomly selected data bases and the model acceptance criteria was applied to them. Final models for the engine categories ALL ENGINES, TURBOJETS, and TURBOFANS were developed utilizing 100 percent of the data records in each category.

This research met the objective which was to develop cost estimating relationships that utilized both engine operating parameters and physical engine characteristics. The final CER developed for the engine category ALL ENGINES (Appendix C-34) utilized 16 independent variables of which seven are operating parameters, e.g., X16 (MAXTIT) and X17 (NORMSFC), and nine of which are physical characteristics, e.g., X1 (CR) and X6 (L). The final model developed for the engine category TURBOJETS (Appendix C-35) utilized 14 variables of which seven were operating parameters and seven were physical characteristics. The TURBOFANS cost model (Appendix C-36) contained six variables of which two were operating parameters and four were physical engine characteristics.

Although the research supported the concept of utilizing both engine operating parameters and physical characteristics, only two (ALL ENGINES and TURBOFANS) of the final three models developed met the minimum R^2 criterion (.8000) at the 95 percent confidence level. However, neither of these models satisfied the remaining two

criteria for model validation, i.e., a predictive efficiency of not less than 95 percent and a percent predictability not less than 50 percent.

CONCLUSIONS

In this section the conclusions drawn from the research effort are presented. It must be realized that some of the conclusions discussed are only applicable within the scope of this study. On the other hand, some may have general applicability.

Initial models developed for the ALL ENGINE and the TURBOFAN category had R^2 values of .9003 and 1.0000, respectively. However, the model developed for the TURBOJET category had an R^2 of only .6101, a value less than the minimum acceptable R^2 value of .8000.

Upon analyzing the models developed for the categories of ALL ENGINES and TURBOJETS it was discovered that only one, X16 (MAXTIT), of the 14 variables in the ALL ENGINE model (Appendix C-13) made a significant marginal contribution to the model's R^2 value. None of the 13 variables in the TURBOJETS model (Appendix C-15) made a significant marginal contribution to the model's R^2 value.

The low R^2 value for the TURBOJETS model and the fact that only X16 in the ALL ENGINES model made a significant marginal contribution to either of the models' R^2 values indicated a potential

problem area. It was surmised that the COST variable data and/or the independent variable data for the TURBOJET and/or the ALL ENGINES category was of questionable reliability. Further, it was felt that the R^2 value for a model developed using data from either of the two categories would depend on which segment of the data was used for model development.

Further analysis was performed to determine if what was surmised was true. One engine data record at a time was held out of each data category and then replaced after a model was developed. This procedure was followed until all engines in each category had been held out. Thus, 28 models were developed for ALL ENGINES, 19 for TURBOJETS and 9 for TURBOFANS. The distribution of R^2 values for the TURBOJETS models yielded an expected R^2 value of .8651. However, the values ranged from .7004 to .9996, a spread of .2992, with seven of the nineteen values in the .7000 to .8000 range, two values between .8000 and .9000 and ten values above .9000. No two R^2 values were the same. The distribution for this category indicated that the data was unstable and that the R^2 value was dependent on the particular engine data records used for model construction. The distribution of the nine values for TURBOFANS models revealed an expected R^2 value of .9988 with values ranging from .9910 to 1.0000, a spread of only .009. The turbofan data consistently provided a model R^2 value greater than .9900 irrespective

of the data records used for model construction. The distribution for ALL ENGINES was better than the distribution for TURBOJETS but worse than the distribution for TURBOFANS. The expected R^2 value for ALL ENGINES was .8901 and the values ranged from .7928 to .9185, a spread of .1257 indicating that the R^2 value was dependent on the engine data records used.

An interim conclusion was made that the turbojet data was unstable, not representative of the population, and could not produce a reliable model. Further, the turbofan data appeared very stable and consistent and when the turbofan and turbojet data were combined in the ALL ENGINES category, the turbofan data greatly influenced the R^2 values for models developed.

An extended analysis was performed on each category utilizing the computer program MREG (Appendix D). One thousand models were developed for each category and an average (expected) value determined, in each category, for R^2 percent predictability, and predictive efficiency. The TURBOJETS category exhibited an R^2 value of .6928, with a range of .8655, a percent predictability of 289 percent with a range of 67,608, and a predictive efficiency of 93.32 percent with a range of 0 to 100 percent. The TURBOFANS category displayed an R^2 value of .9986 with a range of .0027, a percent predictability of 54.25 percent with a range of 127 and a predictive efficiency of 87.70 percent with a range of 0 to 100 percent. The ALL ENGINES category had an R^2 value of .9138 with a range of .3593, a percent

predictability of 418 percent with a range of 72,309, and a predictive efficiency of 96.56 percent with a range of 11.11 to 100 percent.

The extended analysis supported the interim conclusion. The turbojet data is faulty and cannot be expected to produce a useful and/or reliable model. Further, the effect of the turbojet data in the ALL ENGINES category has rendered the data base in that category unreliable for the development of valid models.

The TURBOFAN category appears to have potential for reliable model development. However, the very small sample size, nine data points, as shown in Appendix C-2.1, appears to have restricted that capability.

A high R^2 value can be attained by introducing a large number of variables into the model each with insignificant marginal contribution. However, the model may have virtually no predictive capability. This was evidenced by the 14 variables in the initial ALL ENGINES model (Appendix C-13) and the 13 variables in the initial TURBOJETS model (Appendix C-15). Therefore, a high R^2 value does not imply that the model's predictive capability is good.

The most significant discovery in this research is that models that are statistically significant with high R^2 values indicating high explanatory power, i.e., ability to explain variation in the data, can be developed from questionable and possibly unreliable data. The models appear to be acceptable if taken at face value. Further

analysis, however, may indicate they are impractical and of questionable utility as cost estimators. Therefore, a cost estimating model cannot be accepted on the basis of R^2 value alone. One must analyze the data base for reliability and the model for predictive capability.

As indicated in Appendices C-34, C-35, and C-36 physical characteristics constitute at least 50 percent of the variables in the final model constructed for each engine category. In the ALL ENGINES category, seven of the 16 independent variables were physical characteristics, for TURBOJETS seven of 14 were physical characteristics, and for TURBOFANS four of six were physical characteristics. The high percentage of physical characteristic variables in each model indicates the potential for the development of accurate and usable cost estimating models with engine physical characteristics as major cost drivers. In addition, the technique used in this study of utilizing a percentage of the data base for model development and reserving the remainder of the base for model testing permits discovery of potential problems that would not normally be detected. However, the data base must not be diluted to the point where small degrees of freedom limit its usefulness.

RECOMMENDATIONS FOR FURTHER STUDY

As a result of this research there were several areas identified where extended analysis may provide greater insight in the

development of cost estimating relationships. Areas for further cost estimating study are:

First, cost estimating relationships for predicting depot level overhaul costs for jet engine accessories should be investigated. A study in this area would be complimentary to the research conducted by this study.

Second, a jet engine is considered to have three major sections (a compressor section, a combustor section, and a turbine section). A study should be undertaken to investigate the predictive capability of cost estimating relationships developed to estimate depot level overhaul costs relative to each of these three engine sections.

Third, this study utilized, as the dependent variable, cost expressed on a dollar per flight hour basis. Further research should investigate cost expressed in terms of dollars per end item overhauled.

Finally, depot level overhaul cost estimating relationships developed for jet-core engines or by engine sections, should be combined with cost estimating relationships developed for depot level overhaul of engine accessories. The predictive capability of this combination should then be compared to the predictive capability of total engine (core-engine plus engine accessory) depot level overhaul cost estimating relationships presently in use by the Air Force Logistics Command.

The area of cost estimating within the Department of Defense is wide open and many important discoveries are yet to be made.

APPENDIX A

DATA BASE JET ENGINES AND VARIABLES AND ENGINE CATEGORIES INVESTIGATED

APPENDIX A-1

Independent Variables Considered by Present AFLC Jet Engine Overhaul CERs (23:23)

Engine Dry Weight

Turbine Inlet Temperature

Normal Thrust

Maximum Thrust

Specific Fuel Consumption (at normal thrust)

Specific Fuel Consumption (at maximum thrust)

Thrust to Weight Ratio (at normal thrust)

Thrust to Weight Ratio (at maximum thrust)

Engine Age

Average Flight Time Between Overhaul

APPENDIX A-2

List of Independent Variables

Name	Number	Abbreviation
Number of Compressor Rotors	X1	CR
Number of Compressor Blades	X2	CB
Number of Turbine Rotors	X3	TR
Number of Turbine Blades	X4	TB
Number of Burner Cans	X5	BC
Engine Length	X6	L
Engine Diameter	X7	D
Length to Diameter Ratio	X8	LD
Engine Dry Weight	X9	W
Thrust to Weight Ratio	X10	TW
Rated Thrust at Normal Power	X11	NORMT
Rated Thrust at Military Power	X12	MILT
Rated Thrust at Maximum Power	X13	MAXT
Rated Turbine Inlet Temperature at Normal Power	X14	NORMTIT
Rated Turbine Inlet Temperature at Military Power	X15	MILTIT
Rated Turbine Inlet Temperature at Maximum Power	X16	MAXTIT

Name	Number	Abbreviation
Normal Specific Fuel Consumption	X17	NORMSFC
Military Specific Fuel Consumption	X18	MILSFC
Maximum Specific Fuel Consumption	X19	MAXSFC
Cruise RPM	X20	CRPM
Maximum RPM	X21	MRPM
Fan By-pass Ratio	X22	FBR
Engine Pressure Ratio	X23	EPR
Maximum Rated Airflow	X24	AIR
Average Flight Hours Between Overhaul	X25	AFHBO
Total Flight Hours (FY 1974, 1975, 1976)	X26	FH
Average Flight Hours Per Year	X27	AFH

APPENDIX A-3

Thesis Data Population

J33-A-35	J60-P-5A	TF30-P-3
J57-P-13	J65-W-5	TF30-P-7
J57-P-19	J69-T-25	TF30-P-9
J57-P-21	J75-P-17	TF30-P-100
J57-P-23	J75-P-19/19W	TF33-P-3
J57-P-29	J79-GE-15	TF33-P-5
J57-P-43	J79-GE-17	TF33-P-7
J57-P-55	J85-GE-5	TF33-P-9
J57-P-59	J85-GE-13	TF39-GE-1
J60-P-3		

APPENDIX A-4

Initial Engine Categories Investigated

Category	Number of Engines Per Category 100% of Sample	80% of Sample
1. ALL ENGINES	28	22
2. TURBOJETS	19	15
3. TURBOFANS	9	7
4. TURBOFANS (with afterburner)	4	3
5. TURBOFANS (without afterburner)	5	4
6. TURBOJETS (with afterburner)	10	8
7. TURBOJETS (without afterburner)	9	7

APPENDIX B

REGRESSION ANALYSIS

APPENDIX B

REGRESSION ANALYSIS*

SIMPLE LINEAR REGRESSION

Simple linear regression (SLR) involves regressing a dependent variable (usually plotted on the Y axis) on an independent variable (usually plotted on the X axis). The SLR technique is used to develop a linear regression equation which fits a line to the sample or population data points such that the line can be used as an estimate of Y for a given value of X. The regression line equation is of the form $\hat{Y} = a + bX$ where $\hat{Y}$ is an estimate of the true Y value. Since the regression line will usually not fit all of the data points exactly, there will be a difference between the $\hat{Y}$ and the true value of Y for some values of X. That difference is called the error term or the residual term e_i. Thus, the population regression model equation is expressed as $Y = A + BX_i + e_i$.

<u>Assumptions</u>

The assumptions that are made when working with SLR analysis are:

---

*A working knowledge of regression analysis on the part of the reader is assumed. For a more in-depth presentation of regression analysis, the reader may wish to consult <u>Applied Linear Statistical Models</u> by Neter and Wasserman (14).

1. The error terms (e_i) are independent for all values of X.

2. The expected value of the error terms is zero.

3. The variance of the error terms is constant for all values of X.

4. The error terms for all values of X are distributed normally with a mean of zero and a variance of σ^2.

LEAST SQUARES METHOD

For a regression model to fit the data as well as possible the effect of the error terms must be minimized. Since the (e_i's) may be positive or negative, their sum may equal zero while large e_i values may still exist. The absolute value function of the error terms is not differentiable over the entire range of the function thus a minimum value cannot be determined. The Least Squares Method (LSM) attempts to minimize the sum of the square of the error terms (Σe_i^2). Partial derivatives of the sum are taken with respect to A and B and set equal to zero. Solving the system of normal equations that results yields equations for the estimation of A and B via sample estimators, a and b:

$$\hat{A} = a = \frac{\Sigma x_i y_i - b \Sigma x_i^2}{\Sigma x_i}$$

$$\hat{B} = b = \frac{n \Sigma x_i y_i - \Sigma x_i y_i}{n \Sigma x_i^2 - (\Sigma x_i)^2}$$

Once a and b are computed, the regression equation Y(est) = a + bx is determined for a given value of x.

Since multiple linear regression (MLR) is the technique used in this thesis, SLR will not be discussed further. All statistical tests and acceptance criteria will be explained using MLR terminology and notation.

MULTIPLE LINEAR REGRESSION

Multiple Linear Regression enables regression of Y on two or more independent variables (x_i's). The general MLR population model is:

$$Y = B_o + B_1 x_{i1} + B_2 x_{i2} + \ldots + B_{p-1} x_{i,p-1} + e_i \qquad i = 1, 2, \ldots, n$$

With MLR, p regression coefficients (B's) are estimated using matrix operations and thus the model can be written:

$$Y = XB + e$$

where

$$\underset{(n \times 1)}{Y} = \begin{bmatrix} Y_{31} \\ Y_{32} \\ \cdot \\ \cdot \\ \cdot \\ Y_{3n} \end{bmatrix} \qquad \underset{(n \times p)}{X} = \begin{bmatrix} 1, X_{11}, X_{12}, \ldots, X_{1,p-1} \\ 1, X_{21}, X_{22}, \ldots, X_{2,p-1} \\ \cdot \qquad\qquad\qquad \cdot \\ \cdot \qquad\qquad\qquad \cdot \\ \cdot \qquad\qquad\qquad \cdot \\ 1, X_{n1}, X_{n2}, \ldots, X_{n,p-1} \end{bmatrix}$$

$$\underset{(p\times1)}{B} = \begin{bmatrix} B_{30} \\ B_{31} \\ \cdot \\ \cdot \\ \cdot \\ B_{p-1} \end{bmatrix} \qquad e = \begin{bmatrix} e_{31} \\ e_{32} \\ \cdot \\ \cdot \\ \cdot \\ e_{3n} \end{bmatrix}$$

n = number of sample observations

The Least Squares Method (LSM) is again used to estimate the B's and:

$$B\ (est) = \underset{(p\times1)}{b} = \begin{bmatrix} b_{30} \\ b_{31} \\ \cdot \\ \cdot \\ \cdot \\ b_{p-1} \end{bmatrix} = (X^T X)^{-1}\,(X^T Y)$$

Assumptions

By the Gauss-Markov Theorem, LSM estimators are the best linear unbiased estimators of the population regression coefficients. However, the following assumptions must be valid for the least squares method of MLR to be applicable (5:528-530; 14:38):

1. The four assumptions listed for SLR.
2. A linear model is appropriate.
3. The number of sample observations, n, exceeds the number of parameters to be estimated, p, i.e., $n > p$.
4. The rank of the X matrix is p.
5. The X_i's are observed without error.

Decomposition of Variation

If the assumption that a linear model is appropriate is not made, i.e., no regression of Y on the Xi's, then the best point estimator of Y is its mean ($\overline{Y}$) and there will be variation of the actual Y values about $\overline{Y}$. When Y is regressed on the Xi's, an attempt is made to explain as much of the variation as possible by the regression model.

The total variation (TV) or sum of squares total (SST) may be decomposed into explained variation (EV) or regression sum of squares (SSR) and unexplained variation (UV) or error sum of squares (SSE) such that:

$$TV = EV + UV$$

OR

$$SST = SSR + SSE$$

SSR is the component of total variation explained by the regression model and SSE is the component of total variation unexplained by the regression model. The regression mean square (MSR) is SSR divided by p-1 and the error mean square (MSE) is SSE divided by n-p. These statistical parameters are discussed later in the text. Table I lists computational equations for the respective components of variation.

Table I
Computational Equations for the Components of Variation

Source	Sum of Squares	df	Mean Squares
Explained by regression	$SSR = b^T(X^TY) - n\hat{Y}^2$	p-1	$MSR = \frac{SSR}{p-1}$
Unexplained by regression	$SSE = Y^TY - b^T(X^TY)$	n-p	$MSE = \frac{SSE}{n-p}$
Total	$SST = Y^TY - n\hat{Y}^2$	n-1	

The Coefficient of Determination (R^2)

The greater the portion of total variation explained by the regression model, the better the model fits the data.

The coefficient of multiple determination (R^2) is used as a measure of how well the model fits the data. R^2 is defined as:

$$R^2 = \frac{SSR}{SST}$$

The closer R^2 is to one, the greater the amount of variation explained by the regression model.

F Tests

To test the statistical significance of the overall regression model an F ratio is computed from sample data and compared to a critical F value in an attempt to reduce or resolve the hypotheses test:

$$H_o\text{: All } B_i = 0$$

H_1: At least one $B_i \neq 0$ $i = 1, \ldots, p-1$

The sample F ratio:

$$F = \frac{MSR}{MSE}$$

yields an $F_{p-1, n-p}$ which is compared to $F_{(crit)1-\alpha;\, p-1, n-p}$. F_{crit} is obtained from a standard F-table for a level of significance $1-\alpha$ and for degrees of freedom $p-1, n-p$. If $F < F_{crit}$, H_o cannot be rejected and there is insufficient evidence to conclude H_1: At least one $B_i \neq 0$ $i = 1, \ldots, p-1$. If $F > F_{crit}$, H_o can be rejected and the conclusion made that the overall regression model is statistically significant at the chosen level.

To test the statistical significance of the marginal contribution of the X_i's, the estimated variance-covariance matrix of the b's must be computed:

$$s^2(b) = MSE(X^T X)^{-1}$$

$s(b_i)$ is obtained from the $s^2(b)$ matrix and a t-statistic is computed for the regression coefficient of the respective X_i:

$$t = \frac{b_i}{s(b_i)}$$

The t-statistic is compared to a critical t value, $t_{(crit)\, \alpha/2; n-p}$ obtained from a standard t-table for a level of significance, α and associated

degrees of freedom n-p.

When testing the X_i's simultaneously, each test must be conducted at an equivalent α level so that the overall system of tests maintains a confidence of $(1-\alpha) \times 100\%$. A conservative approach to simultaneous testing which insures at least a $(1-\alpha) \times 100\%$ level of confidence is the Bonferroni technique. This technique yields an equivalent $\alpha(\alpha')$:

$$\alpha' = \frac{\alpha}{p-1}$$

The comparison of t-statistics is used to reduce or resolve the hypothesis test:

$$H_o\text{: } B_i = 0$$

$$H_1\text{: } B_i \neq 0$$

If $|t| > t_{crit}$, reject H_o and conclude that the marginal contribution of X_i is statistically significant at the chosen α level. If $|t| < t_{crit}$, H_o cannot be rejected and there is insufficient evidence to conclude that the marginal contribution is significant (14:146-149).

Confidence Intervals

Since regression models provide only an estimate of the actual value of the dependent variable Y, most of the time the actual value of Y will not coincide with the estimated value, i.e., the actual value will not normally be on the regression surface.

To determine the confidence that may be placed in a regression model as a predictor, confidence intervals can be constructed about the estimates rendered by the model.

A confidence interval (CI) is constructed about the estimates of a regression model at a chosen level of significance, α. A decision maker can then be $(1-\alpha) \times 100\%$ confident that the actual value of Y will be no farther than one half of the CI away from the estimate for Y.

Standard Error of the Estimate

The standard error of the estimate is a measure of dispersion of the error terms (e_i's) about the locus of the regression model estimates, i.e., the regression surface, and is computed:

$$s_{y\,.\,x_1,\ldots,x_{p-1}} = \sqrt{\frac{UV}{n-p}}$$

where:

n = the number of cases in the sample

p = the number of x_i's in the model plus one

UV = the amount of variation unexplained by the model.

APPENDIX C

SUPPORTING DATA AND FORMULATED COST ESTIMATING RELATIONSHIPS

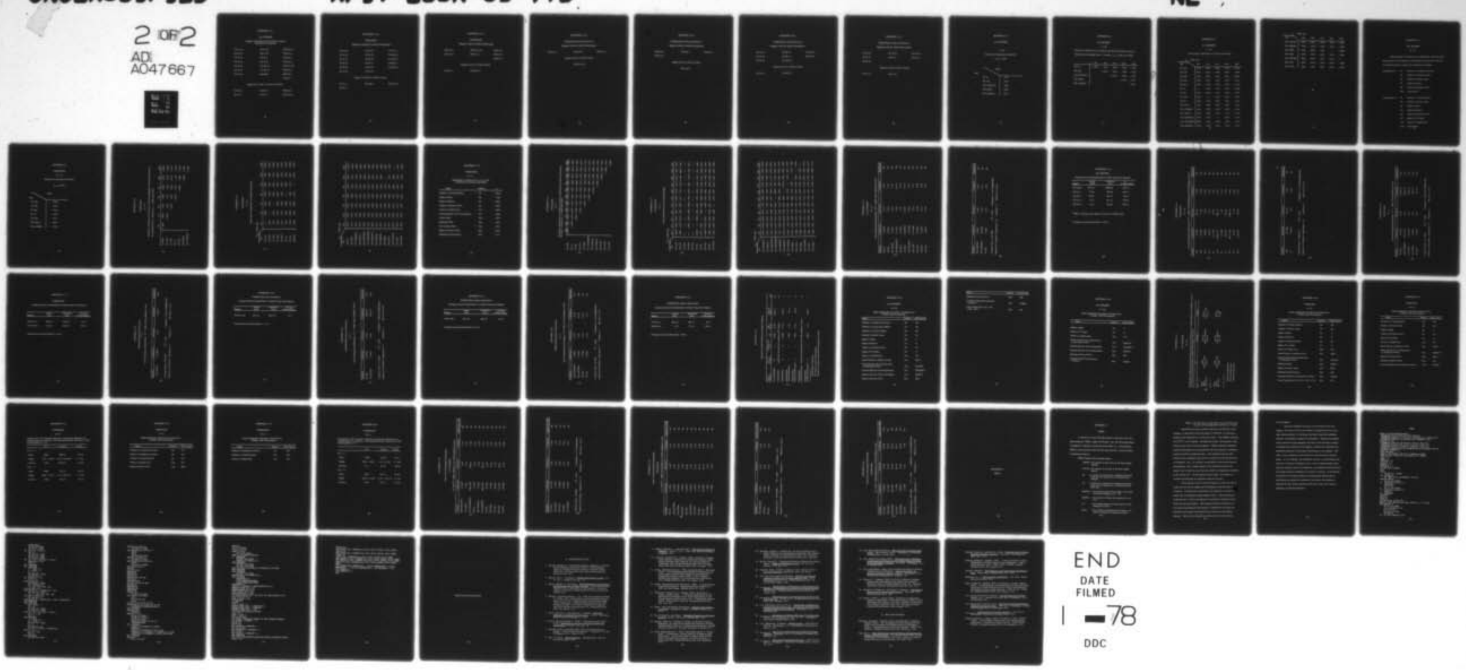
APPENDIX C-1

Independent Variables Considered in the Development of CERs by Engine Category

	Engine Category						
	All Engines	All Turbojets	All Turbofans	Turbofans (W/AB)	Turbofans (W/O AB)	Turbojets (W/AB)	Turbojets (W/O AB)
X1	X	X	X	X	X	X	X
X2	X	X	X	X	X	X	X
X3	X	X	X	X	X	X	X
X4	X	X	X	X	X	X	X
X5	X	X	X	X	X	X	X
X6	X	X	X	X	X	X	X
X7	X	X	X	X	X	X	X
X8	X	X	X	X	X	X	X
X9	X	X	X	X	X	X	X
X10	X	X	X	X	X	X	X
X11	X	X	X	X	X	X	X
X12	X	X	X	X	X	X	X
X13				X		X	
X14							
X15							
X16	X	X	X	X	X	X	X
X17	X	X	X	X	X	X	X
X18	X	X	X	X	X	X	X
X19				X		X	
X20	X	X	X	X	X	X	X
X21	X	X	X	X	X	X	X
X22			X	X	X		
X23	X	X	X	X	X	X	X
X24	X	X	X	X	X	X	X
X25	X	X	X	X	X	X	X
X26	X	X	X	X	X	X	X
X27	X	X	X	X	X	X	X

AD-A047 667 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCHO--ETC F/G 21/5
DEVELOPMENT OF COST ESTIMATING RELATIONSHIPS FOR AIRCRAFT JET C--ETC(U)
SEP 77 R A BREGLIO, R F WRIGHT
UNCLASSIFIED AFIT-LSSR-31-77B NL

2 OF 2
AD
A047667



END
DATE
FILMED
1 -78
DDC

APPENDIX C-2.1

ALL ENGINES

Engines Considered in Correlation Analysis and Model Development

J33-A-35	J57-P-59	J85-GE-13
J57-P-13	J60-P-5A	TF30-P-3
J57-P-19	J65-W-5	TF30-P-7
J57-P-21	J69-T-25	TF30-P-9
J57-P-29	J79-GE-15	TF30-P-100
J57-P-43	J79-GE-17	TF30-P-3
J57-P-55	J85-GE-5	TF33-P-5
		TF33-P-7

Engines Set Aside for Model Validation

J57-P-23	J75-P-17	TF33-P-9
J60-P-3	J75-P-19	TF39-GE-1

APPENDIX C-2.2

TURBOJETS

Engines Considered in Model Development

J33-A-35	J57-P-55	J75-P-19
J57-P-13	J57-P-59	J79-GE-15
J57-P-21	J60-P-5A	J79-GE-17
J57-P-23	J65-W-5	J85-GE-5
J57-P-29	J69-T-25	J85-GE-13
J57-P-43	J75-P-17	

Engines Utilized for Model Testing

J57-P-9	J85-GE-5	G85-GE-13
J60-P-3		

APPENDIX C-2.3

TURBOFANS

Engines Used for Model Development

TF30-P-3	TF30-P-100	TF33-P-5
TF30-P-9	TF33-P-3	TF33-P-7
		TF33-P-9

Engines Used for Model Testing

TF30-P-7	TF39-GE-1

APPENDIX C-2.4

TURBOFANS (with afterburner)

Engines Used for Model Development

TF30-P-3	TF30-P-7	TF30-P-9

Engine Used for Model Testing

TF30-P-100

APPENDIX C-2.5

TURBOFANS (without afterburner)

Engines Used for Model Development

TF33-P-3	TF33-P-7	TF33-P-9
TF33-P-5		

Engine Used for Model Testing

TF39-GE-1

APPENDIX C-2.6

TURBOJETS (with afterburner)

Engines Used for Model Development

J57-P-13	J75-P-17	J79-GE-17
J57-P-23	J75-P-19	J85-GE-5
J57-P-55	J79-GE-15	

Engines Used for Model Testing

J57-P-21	J85-GE-13

APPENDIX C-2.7

TURBOJETS (without afterburner)

Engines Used for Model Development

J33-A-35	J57-P-59	J60-P-5A
J57-P-29	J60-P-3	J65-W-5
J57-P-43		

Engines Used for Model Testing

J57-P-19	J69-T-25

APPENDIX C-3

ALL ENGINES

n = 22

Bivariate Correlations with COST

$r_o \geq \pm .5000$

with	COST
X6 (L)	.6721
X8 (LD)	.5235
X16 (MAXTIT)	.5484
X23 (EPR)	.5847
X25 (AFHBO)	-.6278

APPENDIX C-4

ALL ENGINES

n = 22

Bivariate coefficient of correlation between independent variables which met the modified Rule 1 criterion ($r_o \geq .5000$) with COST.

	X6	X8	X16	X23	X25
X6 (L)	1.0000	.7041	.6492	.7866	-.6142
X8 (LD)		1.0000	.4469	.3229	-.5693
X16 (MAXTIT)			1.0000	.8222	-.6526
X23 (EPR)				1.0000	-.4856
X25 (AFHBO)					1.0000

APPENDIX C-5

ALL ENGINES

n = 66

First Order Coefficients of Partial Correlation

Controlling for \ COST with	X6	X8	X16	X23	X25
X1 (CR)	.6130	.4153	.4329	.4681	-.5686
X2 (CB)	.6525	.4360	.4768	.5199	-.5863
X3 (TR)	.6101	.4865	.4242	.6557	-.5688
X4 (TB)	.6789	.5330	.5238	.7058	-.6116
X5 (BC)	.6667	.5294	.5568	.5636	-.6316
X6 (L)	---	.0956	.1990	.1227	-.3680
X7 (D)	.6716	.5899	.5191	.6226	-.6157
X8 (LD)	.5016	---	.4125	.5155	-.4709
X9 (W)	.6854	.4712	.5003	.5091	-.6005
X10 (TW)	.6907	.4898	.5157	.5645	-.6053
X11 (NORMT)	.6645	.5532	.5143	.6696	-.6152
X12 (MILT)	.6700	.5468	.5145	.6797	-.6168
X16 (MAXTIT)	.4969	.3722	---	.2812	-.4261
X17 (NORMSFC)	.6455	.5366	.5004	.6592	-.5932
X18 (MILSFC)	.6506	.5448	.5035	.6612	-.6014

Controlling for \ COST with	X6	X8	X16	X23	X25
X20 (CRFM)	.7093	.5247	.5630	.5926	-.6285
X21 (MRPM)	.7094	.5243	.5629	.5920	-.6285
X23 (EPR)	.4235	.4359	.1464	---	-.4849
X24 (AIR)	.6883	.5558	.5591	.7321	-.6269
X25 (AFHBO)	.4664	.2595	.2351	.4113	---
X26 (FH)	.6463	.5118	.5307	.5750	-.5948
X27 (AFH)	.6463	.5118	.5307	.5750	-.5948

APPENDIX C-6

ALL ENGINES

N = 22

Representative combinations of independent variables which may produce cost estimating relationships which are least affected by multicollinearity between the independent variables.

Combination #1	X1	Number of Compressor Rotors
	X4	Number of Turbine Blades
	X5	Number of Burner Cans
	X7	Engine Diameter
	X8	Length to Diameter Ratio
	X20	Cruise RPM
Combination #2	X4	Number of Turbine Blades
	X5	Number of Burner Cans
	X6	Engine Length
	X7	Engine Diameter
	X8	Length to Diameter Ratio
	X9	Engine Dry Weight
	X10	Thrust to Weight Ratio
	X20	Cruise RPM

APPENDIX C-7

TURBOJETS

N = 15

Bivariate Correlation with COST

$r_o \geq \pm .5000$

with	COST
X1 (CR)	.5278
X2 (CB)	.5246
X3 (TR)	.5168
X6 (L)	.5705
X8 (LD)	.6313
X23 (EPR)	.5267
X25 (AFHBO)	-.5799

APPENDIX C-8

TURBOJETS

N = 15

Bivariate coefficient of correlation between independent variables which met the modified Rule 1 criterion ($r_o \geq .5000$) with COST.

	X1	X2	X3	X6	X8	X23	X25
X1 (CR)	1.0000	.9804	.9729	.7788	.8156	.9352	-.4662
X2 (CB)		1.0000	.9393	.7133	.7718	.9116	-.4168
X3 (TR)			1.0000	.8384	.8667	.9714	-.4167
X6 (L)				1.0000	.9592	. 383	-.5996
X8 (LD)					1.0000	.8658	-.6700
X23 (EPR)						1.0000	-.3843
X25 (AFHBO)							1.0000

APPENDIX C-9

TURBOJETS

N = 15

First Order Coefficients of Partial Correlation

Controlling for \ COST with	X1	X2	X3	X6	X8	X23	X25
X1 (CR)	---	.0429	.0168	.2993	.4085	.1100	-.4442
X2 (CB)	.0802	---	.0822	.3290	.4181	.1383	-.4667
X3 (TR)	.1263	.1334	---	.2940	.4293	.1213	-.4684
X4 (TB)	.4243	.3976	.4298	.4689	.5597	.4023	-.5261
X5 (BC)	.5199	.5152	.5040	.5644	.6256	.5160	-.5779
X6 (L)	.1621	.2045	.0860	---	.3618	.1081	-.3618
X7 (D)	.5029	.5006	.4913	.5934	.6193	.5035	-.5654
X8 (LD)	.0289	.0759	-.0784	-.1595	---	-.0512	-.2725

Controlling for \ COST with	X1	X2	X3	X6	X8	X23	X25
X9 (W)	.3104	.3267	.2794	.4786	.5199	.3031	-.4719
X10 (TW)	.5315	.5315	.5182	.5862	.6157	.5215	-.5524
X11 (NORMT)	.3530	.3633	.3316	.4469	.5327	.3534	-.4862
X12 (MILT)	.3577	.3638	.3391	.4597	.5393	.3620	-.4924
X16 (MAXTIT)	.2896	.2776	.2887	.4532	.4843	.2902	-.4668
X17 (NORMSFC)	.3214	.2620	.2693	.3744	.4756	.2658	-.4855
X18 (MILSFC)	.4092	.3251	.3868	.3983	.4997	.3271	-.5150
X20 (CRPM)	.3571	.3538	.3447	.4292	.5224	.3577	-.4835
X21 (MRPM)	.3539	.3506	.3412	.4265	.5205	.3543	-.4836
X23 (EPR)	.1171	.1273	.0256	.2783	.4120	---	-.4810
X24 (AIR)	.3674	.3734	.3506	.4860	.5512	.3714	-.5007
X25 (AFHBO)	.3572	.3820	.3715	.3417	.4013	.4040	---
X26 (FH)	.5285	.5332	.5193	.5631	.6247	.5425	-.6041
X27 (AFH)	.5285	.5332	.5193	.5631	.6247	.5425	-.6041

APPENDIX C-10

TURBOFANS

N = 7

Independent Variables with $r_o \geq \pm .7000$
Bivariate Correlation with COST

Name	Number	r_o
Number of Turbine Blades	X4	-.8998
Engine Length	X6	.7398
Engine Diameter	X7	-.8054
Length to Diameter Ratio	X8	.7403
Thrust to Weight Ratio	X10	.8197
Normal Specific Fuel Consumption	X17	.7808
Cruise RPM	X20	.7599
Maximum RPM	X21	.7479
Fan By-pass Ratio	X22	.7071
Engine Pressure Ratio	X23	.8254
Maximum Rated Airflow	X24	.7137

APPENDIX C-11

TURBOFANS

N = 7

Bivariate coefficients of correlation between independent variables which met the Rule 1 criteria ($r_o \geq .7000$) with COST.

	X4	X6	X7	X8	X10	X17	X20	X21	X22	X23	X24
X4 (TB)	1.0000	-.9189	.9544	-.9191	-.9409	-.9643	-.9353	-.9307	.9246	-.9446	.8884
X6 (L)		1.0000	-.9755	1.0000	.7653	.9824	.9957	.9968	-.8790	.8434	-.9864
X7 (D)			1.0000	-.9766	-.8366	-.9765	-.9912	-.9880	.8665	-.8359	.9801
X8 (LD)				1.0000	.7658	.9819	.9962	.9973	-.8770	.8415	-.9875
X10 (TW)					1.0000	.8637	.7975	.7928	-.9088	.9411	-.7252
X17 (NORMSFC)						1.0000	.9853	.9850	-.9462	.9146	-.9537
X20 (CRPM)							1.0000	.9996	-.8789	.8403	-.9909
X21 (MRPM)								1.0000	-.8803	.8398	-.9910
X22 (FBR)									1.0000	-.9648	.8105
X23 (EPR)										1.0000	-.7626
X24 (AIR)											1.0000

APPENDIX C-12

TURBOFANS

N = 7

First Order Partial Correlation Results

Controlling for \ COST with	X4	X6	X7	X8	X10	X17	X20	X21	X22	X23	X24
X1 (CR)	-.8669	.6069	-.7124	.6072	.7434	.6820	.6379	.6174	-.5479	.7999	-.5655
X2 (CB)	-.8862	.7139	-.7758	.7951	.7598	.7277	.7150	.7150	-.7120	.8445	-.6710
X3 (TR)	---	---	---	---	---	---	---	---	---	---	---
X4 (TB)	---	-.5059	.4096	-.5048	-.1838	-.7523	-.5612	-.5612	.7518	-.1715	.4276
X5 (BC)	---	---	---	---	---	---	---	---	---	---	---
X6 (L)	-.8290	---	.0800	-.1967	.4292	.3735	.1936	.1936	-.1771	.5571	.1449
X7 (D)	-.7411	-.3522	---	.4488	-.0450	-.4895	-.5229	-.5229	-.0311	.4676	.6431
X8 (LD)	-.8284	-.0703	-.5701	---	.5843	.4234	.3821	.1936	-.1792	.5573	.1635
X9 (W)	-.8661	.6544	-.7341	.6535	.7605	.7140	.6723	.6560	-.6795	.8238	-.6007
X10 (TW)	-.6634	.3054	-.3816	.3058	---	.2525	.3077	.2810	.1575	.3160	-.3027

Controlling for \ COST with	X4	X6	X7	X8	X10	X17	X20	X21	X22	X23	X24
X11 (NORMT)	-.8753	.6914	-.8274	.6967	.7599	.7069	.7466	.7212	-.5860	.7644	-.7543
X12 (MILT)	.8593	.6468	-.7764	.6515	.7580	.6702	.6954	.6663	-.5628	.7500	-.6741
X16 (MAXTIT)	-.9095	.5004	-.6442	.4982	.6425	.7377	.5440	.5051	-.3171	.6919	-.3443
X17 (NORMSFC)	-.8879	-.2327	-.3193	-.2231	.4610	---	-.0875	.1965	.1565	.4405	.1641
X17 (MILSFC)	-.9771	.4490	-.7458	.4507	.6319	.8891	.5801	.5220	-.2853	.6553	-.3135
X20 (CRPM)	-.8219	-.2794	-.6064	-.2938	.5444	.2881	---	-.6578	-.1265	.5301	.4487
X21 (MRPM)	-.8392	-.1067	-.6487	-.1132	.5601	.3852	.6754	---	-.1549	.5475	.3079
X22 (FBR)	-.9135	-.3508	-.5459	.3536	.5996	.4882	.4105	.3738	---	.7695	-.3396
X23 (EPR)	-.6485	-.1440	-.3727	.1499	.2697	.1135	.2169	.1784	.6001	---	-.2309
X24 (AIR)	-.8264	-.3109	-.7616	.3213	.6260	.4750	.5587	.4322	-.3136	.6204	---
X25 (AFHBO)	-.8392	.5329	-.6671	.5339	.7086	.6159	.5761	.5521	-.4853	.7026	-.4957
X26 (FH)	-.8841	.6903	-.7838	.6907	.7877	.7399	.7151	.6994	-.6707	.8282	-.6577
X27 (AFH)	-.8841	.6903	-.7838	.6907	.7877	.7399	.7151	.6994	-.6707	.8282	-.6577

APPENDIX C-13

ALL ENGINES

N = 22

Cost Estimating Relationship: $R^2 \geq .8000$; 95% Confidence Level

Variable	Coefficient	Std Error	F-Statistic	Significant
X6 (L)	0.9122	0.37	5.79	No
X9 (W)	-0.0611	0.02	6.12	No
X23 (EPR)	4.3241	4.21	1.05	No
X16 (MAXTIT)	-0.3111	0.08	14.75	Yes
X5 (BC)	-1.6966	1.59	1.13	No
X3 (TR)	-43.86938	45.05	0.94	No
X25 (AFHBO)	-0.01376	-0.66	3.54	No
X10 (TW)	-12.71246	9.53	1.77	No
X1 (CR)	16.6924	6.88	5.88	No
X4 (TB)	-0.1623	0.17	0.90	No
X2 (CB)	0.0353	0.09	0.15	No

Variable	Coefficient	Std Error	F-Statistic	Significant
X20 (CRPM)	0.0161	0.00	3.69	No
X11 (NORMT)	0.01052	0.00	2.91	No
X7 (D)	3.2241	2.36	1.86	No

Constant Term: 139.5469 R^2 = .9003   F-Statistic = 4.51 (Significant)

Degrees of Freedom: Regression = 14   Residual = 7   Std Error = 10.66

APPENDIX C-14

ALL ENGINES

Average Percent Predictability of Model Using Test Engines

Engine	Actual Cost*	Predicted Cost	% Predictability
J57-P-23	$20.78	$488.68	2251.0
J60-P-3	11.94	408.84	2237.0
J75-P-17	34.36	444.77	1194.0
J75-P-19	9.61	461.91	4706.0
TF33-P-9	4.61	554.29	11923.0

*COST = Average core-engine overhaul cost/flight hour

Average percent predictability = 4461.0

APPENDIX C-15

TURBOJETS

N = 15

Cost Estimating Relationship: $R^2 < .8000$; Not Significant at 95% Confidence Level

Variable	Coefficient	Std Error	F-Statistic	Significant
X8	6.9033	274.33	0.00	No
X25	-0.0269	0.06	0.18	No
X26	0.0000	0.00	0.00	No
X2	0.0514	0.34	0.02	No
X1	-11.7195	46.74	0.06	No
X18	-3418.2114	7987.70	0.18	No
X4	-0.5289	1.45	0.13	No
X3	-34.6371	370.40	0.00	No
X17	2421.5369	8493.24	0.08	No
X6	0.2477	5.49	0.00	No
X23	-12.9041	88.30	0.02	No

Variable	Coefficient	Std Error	F-Statistic	Significant
X10	18.7888	145.76	0.01	No
X5	-0.8424	7.57	0.01	No

Constant Term: 1308.4733 R^2 = .6101 F-Statistic = 0.12 (Not Significant)

Degrees of Freedom: Regression = 13 Residual = 7 Std error = 33.14

APPENDIX C-16

TURBOFANS

N = 7

Cost Estimating Relationship: $R^2 \geq .8000$; 95% Confidence Level

Variable	Coefficient	Std Error	F-Statistic	Significant
X4 (TB)	-1.6879	0.00	74499.07	Yes
X18 (MILSFC)	-441.9550	2.19	40501.50	Yes
X10 (TW)	-15.4903	0.23	4416.10	Yes
X27 (AFH)	-0.0000	0.00	1067.70	Yes
X2 (CB)	0.0408	0.00	415.96	Yes

Constant Term: 993.2826 $R^2 = 1.0000$ F-Statistic: 46150.40 (Significant)

Degrees of Freedom: Regression = 5 Residual = 1 Std error = 0.14

APPENDIX C-17

TURBOFANS

Average Percent Predictability of Model Using Test Engines

Engine	Actual Cost	Predicted Cost	Percent Predictability
TF-30-P-7	$30.64	$18.02	41.0
TF-39-GE-1	132.41	660.86	399.0

Average percent predictability = 220.0

APPENDIX C-18

TURBOFANS (with afterburner)
N = 3

Cost Estimating Relationship: $R^2 \geq .8000$; 95% Confidence Level

Variable	Coefficient	Std Error	F-Statistic	Significant
X23 (EPR)	-37.9000	1.82	431.54	Yes

Constant Term = 692.8366 $R^2 = .9976$ F-Statistic = 431.54 (Significant)

Degrees of Freedom: Regression = 1 Residual = 1 Std error = 1.29

APPENDIX C-19

TURBOFANS (with afterburner)

Average Percent Predictability of Model Using Test Engines

Engine	Actual Cost	Predicted Cost	Percent Predictability
TF30-P-100	$77.55	$133.39	72.0

Average percent predictability = 72.0

APPENDIX C-20

TURBOFANS (without afterburner)

N = 4

Cost Estimating Relationship: $R^2 \geq .8000$; 95% Confidence Level

Variable	Coefficient	Std Error	F-Statistics	Significant
X9 (W)	-0.0216	0.00	3085.13	Yes
X12 (MILT)	0.0047	0.00	1832.41	Yes

Constant Term = 13.2878 $R^2 = .9999$ F-Statistic = 9209.89 (Significant)

Degrees of Freedom: Regression = 2 Residual = 1 Std error = 0.01

APPENDIX C-21

TURBOFANS (without afterburner)

Average Percent Predictability of Model Using Test Engines

Engine	Actual Cost	Predicted Cost	Percent Predictability
TF39-GE-1	$132.41	$49.85	62.35

Average percent predictability = 62.35

APPENDIX C-22

TURBOJETS (without afterburner)

N = 7

Cost Estimating Relationship: $R^2 \geq .8000$; 95 % Confidence Level

Variable	Coefficient	Std Error	F-Statistic	Significant
X16 (MAXTIT)	-0.0036	0.00	0.18	No
X25 (AFHBO)	-0.0060	0.00	54.97	Yes
X8 (LD)	6.7680	1.63	17.06	No
X9 (W)	0.0004	0.00	1.25	No

Constant Term = 6.1602 $R^2 = .9869$ F-Statistic = 37.87 (Significant)

Degrees of Freedom: Regression = 4 Residual = 2 Std error = 1.11

APPENDIX C-23

TURBOJETS (without afterburner)

Average Percent Predictability of Model Using Test Engines

Engine	Actual Cost	Predicted Cost	Percent Predictability
J57-P-21	$54.64	$40.79	25.0
J85-GE-13	4.76	30.71	545.0

Average percent predictability = 285.0

APPENDIX C-24

Summary of CER Statistics by Engine Category

Engine Category	# I.V.[1] in Model	# Significant I.V.	R^2	Std Error	Avg. % Pred.[2]	N[3]
ALL ENGINES	14	1	.9003	10.66	4461.0	22
TURBOJETS	model did not meet the minimum R^2 criteria					15
TURBOFANS	5	5	1.0000	0.14	220.0	7
TURBOFANS (with afterburner)	1	1	.9976	1.29	72.0	3
TURBOFANS (without afterburner)	2	2	.9999	0.01	62.35	4
TURBOJETS (with afterburner)	model not significant at the 95% confidence level					8
TURBOJETS (without afterburner)	4	1	.9869	1.11	285.0	7

[1] Independent Variables

[2] Average Percent Predictability

[3] Number of engine cases used to develop the CER

APPENDIX C-25

ALL ENGINES

N = 22

Initial Independent Variables Considered for Further CER Development

Name	Number	Abbreviation
Number of Compressor Rotors	X1	CR
Number of Compressor Blades	X2	CB
Number of Turbine Blades	X4	TB
Number of Burner Cans	X5	BC
Engine Length	X6	L
Engine Diameter	X7	D
Length to Diameter Ratio	X8	LD
Engine Dry Weight	X9	W
Thrust to Weight Ratio	X10	TW
Rated Thrust at Military Power	X12	MILT
Rated Turbine Inlet Temperature at Maximum Power	X16	MAXTIT
Normal Specific Fuel Consumption	X17	NORMSFC
Military Specific Fuel Consumption	X18	MILSFC
Engine Pressure Ratio	X23	EPR

Name	Number	Abbreviation
Maximum Rated Airflow	X24	AIR
Average Flight Hours Between Overhaul	X25	AFHBO
Total Flight Hours (FY 1974, 1975, 1976)	X26	FH

APPENDIX C-26

ALL ENGINES

N = 22

Final Independent Variables Considered for Further CER Development

Name	Number	Abbreviation
Engine Length	X6	L
Engine Dry Weight	X9	W
Thrust to Weight Ratio	X10	TW
Rated Turbine Inlet Temperature at Maximum Power	X16	MAXTIT
Normal Specific Fuel Consumption	X17	NORMSFC
Military Specific Fuel Consumption	X18	MILSFC
Maximum Rated Airflow	X24	AIR
Average Flight Hours Between Overhaul	X25	AFHBO

APPENDIX C-27

ALL ENGINES

N = 22

Comparison of R^2, percent prediction, and percent efficiency for varying degrees of freedom (D.F.). One thousand random selections of 80% of the sample (N= 22).

	R^2	% Pred.[1]	% Eff[2]
D.F. = 4			
Mean	.9138	417.69	96.56
Range	.6402 - .9995	14.99 - 72324.44	11.11 - 100
Std. Dev.[3]	.0571	2785.59	12.39
D.F. = 7			
Mean	.8224	394.01	94.03
Range	.4325 - .9620	0.57 - 84457.5	0 - 100
Std. Dev.	.0839	3415.79	13.88

[1] Percent Prediction

[2] Percent Efficiency

[3] Standard Deviation

APPENDIX C-28

TURBOJETS

N = 19

Initial Independent Variables Considered for Further CER Development

Name	Number	Abbreviation
Number of Turbine Blades	X4	TB
Number of Burner Cans	X5	BC
Engine Length	X6	L
Engine Diameter	X7	D
Length to Diameter Ratio	X8	LD
Engine Dry Weight	X9	W
Thrust to Weight Ratio	X10	TW
Rated Thrust at Military Power	X12	MILT
Rated Turbine Inlet Temperature at Maximum Power	X16	MAXTIT
Maximum RPM	X21	MRPM
Engine Pressure Ratio	X23	EPR
Maximum Rated Airflow	X24	AIR
Average Flight Hours Between Overhaul	X25	AFHBO
Total Flight Hours (FY 1974, 1975, 1976)	X26	FH

APPENDIX C-29

TURBOJETS

N = 19

Final Independent Variables Considered for Further CER Development

Name	Number	Abbreviation
Number of Turbine Blades	X4	TB
Number of Burner Cans	X5	BC
Engine Length	X6	L
Length to Diameter Ratio	X8	LD
Engine Dry Weight	X9	W
Thrust to Weight Ratio	X10	TW
Rated Thrust at Military Power	X12	MILT
Rated Turbine Inlet Temperature at Maximum Power	X16	MAXTIT
Engine Pressure Ratio	X23	EPR
Maximum Rated Airflow	X24	AIR
Average Flight Hours Between Overhaul	X25	AFHBO

APPENDIX C-30

TURBOJETS

N = 19

Comparison of R^2, percent prediction, and percent efficiency for varying degrees of freedom. One thousand random selections of 90% of the sample (N = 19).

	R^2	% Pred.	% Eff.
D.F. = 4			
Mean	.6928	288.99	93.32
Range	.1333 - .9988	2.18 - 67608.20	0 - 100
Std Dev.	.1668	2229.94	21.32
D.F. = 7			
Mean	.5922	191.21	92.74
Range	.3925 - .9975	1.46 - 5223.63	0 - 100
Std Dev.	.1448	512.69	21.23

APPENDIX C-31

TURBOFANS

N = 9

Initial Independent Variables Considered for Further CER Development

Name	Number	Abbreviation
Number of Compressor Rotors	X1	CR
Number of Compressor Blades	X2	CB
Number of Turbine Blades	X4	TB
Thrust to Weight Ratio	X10	TW
Engine Pressure Ratio	X23	EPR

APPENDIX C-32

TURBOFANS

N = 9

Final Independent Variables Considered for Further CER Development

Name	Number	Abbreviation
Number of Compressor Rotors	X1	CR
Number of Turbine Blades	X4	TB
Thrust to Weight Ratio	X10	TW

APPENDIX C-33

TURBOFANS

N = 9

Comparison of R^2, percent prediction, and percent efficiency for varying degrees of freedom. One thousand random selections of 90% of the sample (N = 7).

	R^2	% Pred.	% Eff.
D.F. = 1			
Mean	.9986	54.25	87.7
Range	.9972 - .9999	.93 - 127.66	0 - 100
Std Dev.	.49	39.41	32.86
D.F. = 4			
Mean	.9015	141.51	86.82
Range	.8183 - 1.0000	6.47 - 1394.21	0 - 100
Std Dev.	.0431	209.71	33.02

APPENDIX C-34

Final Cost Estimating Relationship

ALL ENGINES

N = 28; $R^2 > .8000$; Significant at 95% Confidence Level

Variable	Coefficient	Std Error	F-Statistic	Significant (95%)
X16 (MAXTIT)	-0.1084	0.08	1.51	No
X24 (AIR)	0.3411	0.19	3.01	No
X17 (NORMSFC)	282.5188	528.69	0.28	No
X6 (L)	0.0705	0.64	0.01	No
X25 (AFHBO)	-0.0043	0.00	0.62	No
X18 (MILSFC)	-224.2853	533.20	0.17	No
X9 (W)	-0.0160	0.02	0.45	No
X10 (TW)	-2.7303	15.12	0.03	No
X5 (BC)	-2.0411	2.07	0.96	No
X8 (LD)	12.4547	25.67	0.23	No
X1 (CR)	13.3987	7.67	3.05	No

Variable	Coefficient	Std Error	F-Statistic	Significant (95%)
X2 (CB)	-0.1339	0.08	2.28	No
X23 (EPR)	9.1030	6.44	1.99	No
X26 (FH)	-0.0000	0.00	0.30	No
X4 (TB)	-0.1687	0.17	0.94	No
X12 (MILT)	-0.0071	0.00	0.77	No

Constant Term: 78.2070 $R^2 = .8737$ F-Statistic = 4.75 (Significant)

Degrees of Freedom: Regression = 16 Residual = 11 Std Error = 15.55

APPENDIX C-35

Final Cost Estimating Relationship

TURBOJETS

N = 19; $R^2 < .8000$; Not Significant at 95% Confidence Level

Variable	Coefficient	Std Error	F-Statistic	Significant (95%)
X6 (L)	-0.4070	1.20	0.11	No
X25 (AFHBO)	-0.0066	0.01	0.23	No
X23 (EPR)	-38.6610	51.64	0.56	No
X26 (FH)	0.0000	0.00	1.18	No
X8 (LD)	13.8403	45.02	0.09	No
X12 (MILT)	-0.0329	0.03	0.89	No
X10 (TW)	4.5044	19.54	0.05	No
X5 (BC)	5.5013	10.32	0.28	No
X7 (D)	14.4487	14.45	0.99	No
X24 (AIR)	1.7892	2.10	0.72	No
X21 (MRPM)	0.0096	0.01	0.77	No

Variable	Coefficient	Std Error	F-Statistic	Significant (95%)
X16 (MAXTIT)	0.9627	1.05	0.82	No
X4 (TB)	0.5521	0.64	0.73	No
X9 (W)	0.0064	0.04	0.02	No

Constant Term: -1932.9688 $R^2 = .6404$ F-Statistic = 0.51

Degrees of Freedom: Regression = 14 Residual = 4 Std Error = 16.89

APPENDIX C-36

Final Cost Estimating Relationship

TURBOFANS

N = 9; $R^2 > .8000$; Significant at 95% Confidence Level

Variable	Coefficient	Std Error	F-Statistic	Significant (95%)
X1 (CR)	-53.8662	11.84	20.68	Yes
X10 (TW)	-51.6971	13.56	14.51	No
X22 (FBR)	104.9391	13.20	63.13	Yes
X4 (TB)	-1.1785	00.17	43.57	Yes
X23 (EPR)	20.8672	4.83	18.61	Yes
X2 (CB)	0.0785	0.06	1.28	No

Constant Term: 1037.2892 $R^2 = .9976$ F-Statistic = 141.42 (Significant)

Degrees of Freedom: Regression = 6 Residual = 7 Std Error = 4.30

APPENDIX D

MREG

APPENDIX D

MREG

To perform a more thorough analysis of the data, the computer program, MREG, pages 136 through 140, was developed based on programs utilized in a previous thesis effort (9). The program, MREG, is structured so that the user may choose to run any number of iterations desired.

MREG requires the following inputs:

JROWS = the number of data cases in the data sample matrix

JCOLS = the number of columns in the data sample matrix

T1 = t-value for constructing confidence intervals when the test engines are considered individually

T2 = t-value for constructing confidence intervals when the test engines are considered simultaneously

SAMPLE = the percent of the data samples to be used for model building, e.g., .80.

ITIMES = the number of times the program is to be run

JY = the column index in the data matrix of the dependent variable

JCOL = the number of independent variables to be used in building the regression model

IND(J) = the individual column indices in the data matrix for each of the independent variables to be used

The MREG program randomly splits the data file into a percentage, as specified by the value input for SAMPLE, for building a model and the remainder for testing the model. Then MREG computes the $(X^TX)^{-1}$ or "C" matrix, develops the model, and computes a predicted cost for each of the test engines. MREG computes confidence interval half-widths for each predicted cost and computes a confidence interval around the predicted value. The confidence interval half-widths and the confidence intervals are computed for the desired level of confidence, e.g., 95 percent, as specified by the t-values input to the program. The t-values depend on the difference between the sample size (N) and one more than the number of independent variables in the equation (P). As the difference gets larger, the degrees of freedom increase and the confidence interval narrows.

The predicted cost for each test engine is compared with the actual cost for the test engine and the difference between them is computed. The difference is divided by the predicted cost and the actual cost to determine its percentage of each. The percentage of predicted cost is used as the indicator of predictive capability of the model for each test engine. The average predictive capability for the model considering all test engines is computed by summing the individual percentages and dividing by the number of test engines utilized. This value is retained in a data array for each iteration

of the program.

Once the confidence intervals are developed for the test engines, the actual cost for each engine is compared with the confidence interval limits. If the actual cost falls within the confidence interval, an efficiency counter is incremented. When all test engine actual costs have been compared, the value of the efficiency counter is divided by the number of test engines considered to determine the predictive efficiency of the model considering all test engines. The value is also retained in a data array for each iteration of the program. If, for example, the confidence interval is constructed on the basis of a 95 percent confidence level, it can be expected that in the long run, possibly after 1000 iterations, an acceptable model will have an average predictive efficiency of at least 95 percent. If the degrees of freedom are increased either by increasing the sample size or decreasing the number of variables in the model, the confidence interval will then narrow and the model may or may not be able to maintain a 95 percent efficiency.

MREG

```
*$RUN *=(ULIB)GRADLIB/TSS,R
 PARAMETER P1=33,P2=30,P3=P2/2,P4=2*P1
 DIMENSION CONF(P1,1),EE(1,P1),BUILD(P2,P1),TEST(P2,P1)
 DIMENSION COSTP(1,P2),COSTD(1,P2),COSTPCT(1,P2)
&COSTPCP(1,P2)
 DIMENSION X(P2,P1),XT(P1,P2),Y(P2,1),YT(1,P2)
 DIMENSION HOLD(P1,P4),G(P1,1),B(P1,1),BT(1,P1)
 DIMENSION ST(P2,P1),IND(P1),IDUM(P1),CON(1,P1)
 DIMENSION PARRAY(100,20),EARRAY(100,20),R2ARRAY(100,20)
 INTEGER POINTR(P2)
 REAL MSE,MSR
 READ(5,101)JROWS,JCOLS,T1,T2,SAMPLE,ITIMES
 READ(10,101)((ST(J,K),K=1,JCOLS),J=1,JROWS)
 ICOL = 0
 IROW = 1
 NTSTCTR = 0
 NSOLCTR = 0
 NEGVAL = 0
 NODEGFRE = 0
 DO 999 IM=1,ITIMES
 JB=1
 JTB=1
     DO 20 J=1, JROWS
 101 FORMAT(V)
     IF (RND(-1.0).LE.SAMPLE) GO TO 9
 POINTR(JTB)=J
     DO 12 L=1,JCOLS
 12  TEST(JTB,L)=ST(J,L)
     JTB=JTB+1
     GO TO 20
 9   DO 11 I=1,JCOLS
 11  BUILD(JB,I)=ST(J,I)
     JB=JB+1
 20  CONTINUE
 JTB=JTB-1
 JB=JB-1
 JROW=JB
 IF(JTB.EQ.0) GO TO 996
 READ(11,101,END=978)JY,JCOL,(IND(J),J = 1,JCOL)
 978 REWIND 11
     DO 26 I=1,JROW
     Y(I,1)=BUILD(I,JY)
     X(I,1)=1.0
     DO 26 J=1,JCOL
     JJ=IND(J)
 26  X(I,J+1)=BUILD(I,JJ)
```

```
      JCOL=JCOL+1
      DO 30 K=1,JROW
      DO 30 J=1,JCOL
30    XT(J,K) = X(K,J)
      XX = 0
      DO 70 J=1, JCOL
      DO 60 L=1, JCOL
      DO 50 K=1, JROW
50    XX = XX + XT(J,K) * X(K,L)
      HOLD(J,L) = XX
      XX = 0
60    CONTINUE
70    CONTINUE
      K = JCOL + 1
      L = JCOL * 2
      N = K - 1
      DO 100 M=1, JCOL
      DO 90 J=K, L
90    HOLD(M,J) = 0
      N = N + 1
100   HOLD(M,N) = 1
      DO 160 K=1, JCOL
      XX = HOLD(K,K)
      IF (XX) 110, 998, 110
110   DO 120 J=1,L
120   HOLD(K,J) = HOLD(K,J) / XX
      DO 150 J=1, JCOL
      IF (J - K) 130, 150, 130
130   XXX = HOLD(J,K)
      DO 140 M=1, L
140   HOLD(J,M) = HOLD(J,M) - XXX * HOLD(K,M)
150   CONTINUE
160   CONTINUE
      KK = JCOL + 1
      XX = 0
      DO 190 J=1, JCOL
      DO 180 K=1, JROW
180   XX = XX + XT(J,K) * Y(K,1)
      G(J,1) = XX
      XX = 0
190   CONTINUE
      L = JCOL + 1
      LL = JCOL * 2
      DO 210 J=1, JCOL
      M = 1
      DO 200 K=L, LL
      XX = XX + G(M,1) * HOLD(J,K)
200   M = M + 1
      B(J,1) = XX
      XX = 0
210   CONTINUE
      DO 220 J=1,JROW
```

```
220 YT(1,J)=Y(J,1)
    DO 230 J=1,JROW
230 XX=XX+YT(1,J)*Y(J,1)
    YY=XX
    XX=0
    DO 240 J=1,JCOL
240 BT(1,J)=B(J,1)
    DO 250 J=1,JCOL
250 XX=XX+BT(1,J)*G(J,1)
    ZZ=0
    DO 255 J=1,JROW
255 ZZ=ZZ+Y(J,1)
R=JROW
ZZ=(ZZ*ZZ)/R
SST=YY-ZZ
SSR=XX-ZZ
SSE=YY-XX
RR=(XX-ZZ)/(YY-ZZ)
IR=JCOL-1
IX=JROW-JCOL
IF(IX.LT.1)GO TO 969
S2=SSE/IX
MSR=SSR/IR
MSE=SSE/IX
F=MSR/MSE
IDUM(1)=0
    DO 264 I=1,JCOL
264 IDUM(I+1)=IND(I)
    N=0
    DO 265 K=L,LL
    N=N+1
IF(S2.LT.(0.0))GO TO 975
IF(HOLD(N,K).LT.(0.0))GO TO 976
    XX=SQRT(S2)*SQRT(HOLD(N,K))
265 CONTINUE
JROW=JB+JTB
    SVAR=0.0
PREDICT = 0.0
    DO 88 I=1,JROW
    COSTP(I,1)=B(1,1)
    DO 89 L=2,JCOL
    VAR=ST(I,IND(L-1))*B(L,1)
    SVAR=SVAR+VAR
89  CONTINUE
    COSTP(I,1)=COSTP(I,1)+SVAR
    SVAR=0.0
    COSTD(I,1)=COSTP(I,1)-ST(I,JY)
    COSTPCP(I,1)=(COSTD(I,1)/COSTP(I,1))*100
    COSTPCT(I,1)=(COSTD(I,1)/ST(I,JY))*100
88  CONTINUE
ECTR=0.0
    DO 4 I=1,JTB
```

```
      VAL=0.0
      DO 5 J=1,JCOL
      5 EE(1,J)=0.0
      CON(1,1)=1.0
           DO 6 J=1,JCOL-1
      CON(1,J+1)=TEST(I,IND(J))
      6    CONTINUE
           DO 800 II=1,JCOL
           CONF(II,1)=CON(1,II)
      800  CONTINUE
           DO 620 I2=1,JCOL
           DO 630 I3=1,JCOL
      EE(1,I2)=EE(1,I2)+CON(1,I3)*HOLD(I3,I2+JCOL)
      630  CONTINUE
      620  CONTINUE
           DO 660 I3=1,JCOL
      VAL=VAL+EE(1,I3)*CONF(I3,1)
      660  CONTINUE
           VALU=VAL+1.0
           CI=T1*SQRT(VALU*MSE)
           CIA=T2*SQRT(VALU*MSE)
      IN=POINTR(I)
      PREDICT = PREDICT + ABS(COSTPCP(IN,1))
      RANGE1=COSTP(IN,1)+CI
      RANGE2=COSTP(IN,1)-CI
      RAG1=COSTP(IN,1)+CIA
      RAG2=COSTP(IN,1)-CIA
      IF((RAG1-TEST(I,JY)).LT.(0.0).OR.(RAG2-TEST(I,JY))
     &.GT.(0.0))GO TO 4
      ECTR = ECTR + 1.0
      4 CONTINUE
      TB=JTB
      ICOL = ICOL + 1
      PARRAY(IROW,ICOL) = PREDICT/TB
      EARRAY(IROW,ICOL) = ECTR/TB
      R2ARRAY(IROW,ICOL) = RR
      IF((ICOL-20).NE.0)GO TO 999
      ICOL = 0
      IROW = IROW + 1
      995 GO TO 999
      976 PRINT,"NEGATIVE VALUES IN THE INVERSE MATRIX"
      975 NEGVAL = NEGVAL + 1
      PRINT 977
      GO TO 999
      996 NTSTCTR = NTSTCTR + 1
      997 GO TO 999
      969 NODEGFRE = NODEGFRE + 1
      GO TO 999
      998 NSOLCTR = NSOLCTR + 1
      999 CONTINUE
      992 PRINT 980,NTSTCTR,NSOLCTR,NEGVAL,NODEGFRE,ITIMES
      ICOLS = 20
```

```
 PRINT 979
 991 PRINT 1Ø1,((EARRAY(III,JJJ),JJJ=1,ICOLS),III=1,IROW)
 PRINT 979
 99Ø PRINT 1Ø1,((PARRAY(III,JJJ),JJJ=1,ICOLS),III=1,IROW)
 PRINT 979
 PRINT 1Ø1,((R2ARRAY(III,JJJ),JJJ=1,ICOLS),III=1,IROW)
 994 WRITE(12,1Ø1)((EARRAY(III,JJJ),JJJ=1,ICOLS),III=1,IROW)
 993 WRITE(13,1Ø1)((PARRAY(III,JJJ),JJJ=1,ICOLS),III=1,IROW)
 WRITE(14,1Ø1)((R2ARRAY(III,JJJ),JJJ=1,ICOLS),III=1,IROW)
 STOP
 98Ø FORMAT(1ØX,1ØHNTSTCTR = ,I4,5X,1ØHNSOLCTR = ,I4,5X,
&1ØHNEGVAL =  ,I4//5X,11HNODEGFRE = ,I4,9HITIMES = ,I4)
 977 FORMAT(/)
 979 FORMAT(////)
 END
```

SELECTED BIBLIOGRAPHY

A. REFERENCES CITED

1. Barger, Michael A., and Richard McNally, Engineers, Air Force Aero Propulsion Laboratory (AFAPL), Turbine Engine Division, Turbine Propulsion Branch, Wright-Patterson AFB, OH. Personal interviews conducted intermittently March to June, 1977.

2. Bennet, John J. "Comment," Defense Management Journal, Vol. 12, No. 1, January, 1976, pp. 1-4.

3. Black, Thelma L., and others. Depot Maintenance Cost Estimating Relationships for Fighter Aircraft. Cost Analysis Paper MCCC-70-001. Directorate of Cost Analysis, Air Force Logistics Command, Wright-Patterson Air Force Base OH. March, 1970.

4. Breglio, Lieutenant Robert A., Jr., USN, and Lieutenant Richard F. Wright, USN. "Development of Cost Estimating Relationships for Jet 'Core' Engine Overhaul Cost Using Multiple Linear Regression." Unpublished statistics project LSSR Statistics II Class 77B, School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, March, 1977.

5. Clark, Charles T., and Lawrence L. Schkade. Statistical Analysis for Administrative Decisions. 2nd ed. Cincinnati: South-Western Publishing Co., 1974.

6. Evans, General William J., USAF. "Modernizing and Cutting Costs," Supplement to the Air Force Policy Letter For Commanders, October, 1976, pp. 1-8.

7. Executive Office of the President, Office of Management and Budget. "Major System Acquisitions." Circular No. A-109. Washington, D.C., April 5, 1976.

8. Fox, J. Ronald. Arming America. Massachusetts: Harvard University Press, 1974.

9. Kendall, Maurice G., and Alan Stuart. The Advanced Theory of Statistics. 2nd ed., Vol. 2. New York: Hafner Publishing Company, 1967.

10. Lawrence, Frederick P., Captain, USAF. Instructor of Quantitative Methods, Department of Quantitative Studies (LSGQ), School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH. Personal interviews conducted intermittently April to August, 1977.

11. Lynch, Lieutenant Lynn M., USN, and Captain Neil V. Raymond, USAF. "Cost Estimating Relationships for Predicting Life Cycle Costs of Inertial Measurement Unit Maintenance." Unpublished master's thesis, SLSR 16-75A, School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, August, 1975.

12. Miller, Lieutenant Colonel Edward P., USAF. "Current System Acquisition Realities." Unpublished research report No. 5349, Air War College, Maxwell AFB, AL, 1974.

13. Mullineaux, Rodney W. J., Captain, USAF, and Michael A. Yanke, Captain, USAF. "A Proposed Methodology for the Estimation of Jet Engine Costs in the Early Phases of the Weapon System Acquisition Process." Unpublished master's thesis. SLSR 02-76A, AFIT/SL, Wright-Patterson AFB, OH, June, 1976.

14. Neter, John, and William Wasserman. Applied Linear Statistical Models. Homewood, Illinois: Richard D. Irwin, Inc., 1974.

15. Nie, Norman H., and others. Statistical Package for the Social Sciences, 2nd ed. New York: McGraw-Hill, 1970.

16. Panella, Robert F., Michael A. Barga, and Richard McNally, Engineers, Air Force Aero Propulsion Laboratory (AFAPL), Turbine Engine Division, Turbine Propulsion Branch, Wright-Patterson AFB, OH. Personal interview, 21 March 1977.

17. Ross, Captain Milton C., USAF, and Captain Gerald L. Yarger, USAF. "A Parametric Costing Model for Flight Simulator Acquisition." Unpublished master's thesis, SLSR 22-76B, School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, September, 1976.

18. Steinlage, Roger W., Statistician, Air Force Logistics Command, Directorate of Management and Budget, Cost Analysis Branch (ACRCC), Wright-Patterson AFB, OH. Personal interviews conducted intermittently February to June, 1977.

19. Shorey, Russell R. "Managing Downstream Weapons Acquisition Costs." Defense Management Journal, Vol. 12, No. 1, January, 1976, pp. 10-17.

20. Ulsamer, Edgar. "USAF's Campaign to Cut Logistics Costs," Air Force Magazine, September, 1975, pp. 82-87.

21. U.S. Air Force Logistics Command. Air Force Acquisition Logistics Division (AFLC) Implementation Plan. AFLC Programming Plan 76-17. Wright-Patterson AFB, OH: AFLC/XRX, August, 1976.

22. __________. Considerations in Preparing Logistic Support Cost Estimate: A Generalized Cost Model for an Aircraft System. AFLCP 173-1. Wright-Patterson Air Force Base, OH. 5 April 1968.

23. __________. Estimating Depot Level Maintenance Costs for Air Force Aircraft. AFLCP 173-4. Wright-Patterson Air Force Base, OH. January, 1977.

24. U.S. Department of the Air Force. Department of Defense Procurement Management Review: Aircraft Gas Turbine Engine Acquisition and Logistics Support. Washington, D.C. February, 1976.

25. __________. Department of Defense Deflators FY-76 Base Year. AFR 173-10, Attachment 49, May 16, 1975. Washington: Government Printing Office, 1975.

26. U.S. Department of Defense. Design-to-Cost. DOD Directive 5000.28. Washington, D.C.: Government Printing Office, May 23, 1975.

27. __________. Life Cycle Costing Guide for Systems Acquisition (Interim). Washington, D.C.: Government Printing Office, January, 1973.

28. __________. Major System Acquisition Process. DOD Directive 5000.2. Washington: Government Printing Office, January 18, 1977.

29. U.S. Department of Defense. OSD Cost Analysis Improvement Group. DOD Directive 5000.4. Washington: Government Printing Office, 13 June 1973.

30. U.S. General Accounting Office. The Process for Identifying and Establishing Requirements for Major Weapon Systems in the Department of Defense. B-163058. Washington, D.C., 23 October 1974.

31. U.S. Government: Office of the Assistant Secretary of Defense (Installations and Logistics). Life Cycle Costing for Systems Acquisition (Interim, LCC-3). Washington, D.C.: Government Printing Office, January, 1973.

32. Watson, F., Captain, USAF, Air Force Logistics Command, Directorate of Management and Budget, Cost Analysis Division (ACRCC), Wright-Patterson AFB, OH. Personal interviews conducted intermittently May and June, 1977.

33. Wonnacott, Thomas H., and Ronald J. Wonnacott. Introductory Statistics for Business and Economics. New York: John Wiley and Sons, Inc., 1972.

34. Zambo, Leslie J., Major, USAF. Instructor of Quantitative Methods, Department of Quantitative Studies (LSGQ), School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH. Personal interviews conducted intermittently June to August, 1977.

B. RELATED SOURCES

Davidson, George H., Captain, USAF, and Raymond E. Griffiths, Captain, USAF. "An Investigation of Cost-Estimating Techniques Used to Estimate Jet Engine Operation and Support Costs." Unpublished master's thesis proposal. SLSR 1-77A, AFIT/SL, Wright-Patterson AFB, OH, December, 1976.

Dun, John L. Depot Maintenance Cost-Estimating Relationships for Turbine Powered Aircraft. Cost Analysis Paper MCC-68-003. Directorate of Cost Analysis, Air Force Logistics Command, Wright-Patterson Air Force Base, OH, June, 1968.

Heward, James H., and Peter M. Steele. Business Control through Multiple Regression Analysis. New York: John Wiley and Sons, Inc., 1973.

Homich, George H., Captain, USAF. "A Critical Review of Depot Maintenance Overhaul Costs for Aircraft Engines." Unpublished master's thesis. SLSR 3-71B, AFIT/SL, Wright-Patterson AFB, OH, August, 1971.

Ostwald, Phillip F. Cost Estimating for Engineering and Management. New Jersey: Prentice-Hall, Inc., 1974.

Reichman, W. J. Use and Abuse of Statistics. New York: Oxford University Press, 1961.

Stark, Arthur E., Captain, USAF, and James K. Greene, Captain, USAF. "An Investigation of the Relationship of Section Production Costs to Total Production Costs of Gas Turbine Engines." Unpublished master's thesis proposal. SLSR 34-77A. AFIT/SL, Wright-Patterson AFB, OH, January, 1977.

U.S. Air Force Systems Command. Cost Estimating Procedures. AFSCM 173-1. Andrews Air Force Base, Washington, D.C., 28 November 1967.

U.S. Department of the Air Force. USAF Cost Estimating Relationship/Cost Factors Program. AFR 173-2. Washington: Government Printing Office, 16 December 1971.

__________. Independent Cost Analysis Program. AFR 173-11. Washington: Government Printing Office, 1974.

Yarger, Gerald L., Captain, USAF, and Milton C. Ross, Captain, USAF. "A Parametric Costing Model for Flight Simulator Acquisition." Unpublished master's thesis. SLSR 22-76B, AFIT/SL, Wright-Patterson AFB, OH, September, 1976.